Bitcoin Heist Ransonware Analytics

Saipriya Kesoju
CSBS, VNR Vignana Jyothi Institute of engineering and technology (Affiliated to J.N.T.U, Hyderabad)
Bachupally, India
saipriyak0107@gmail.com

Shivani Kongara
CSBS, VNR Vignana Jyothi Institute of engineering and technology (Affiliated to J.N.T.U, Hyderabad)
Bachupally, India
shivaniireddyk@gmail.com

***Abstract*—Ransomware attacks have emerged as one of the most critical threats in the digital era, leveraging cryptocurrency like Bitcoin to demand payments and evade detection. This study addresses the growing challenge of understanding these attacks by analyzing Bitcoinheist ransomware data through Power BI. By focusing on key metrics such as attack frequency, ransom amounts, and Bitcoin address networks, the research uncovers critical patterns shaping the ransomware landscape. Time-series analyses highlight how these threats have evolved, geospatial mapping identifies regional hotspots, and sector-based insights reveal industries under significant risk. Network diagrams further expose connections between Bitcoin addresses, offering a glimpse into attacker strategies. With Power BI's interactive capabilities, users can delve deeper into trends by filtering data by time, ransomware family, or industry, These insights are crucial for equipping cybersecurity experts, policymakers, and law enforcement to counteract ransomware with more targeted and effective defenses.**

Keywords—: Ransomware Threats, Bitcoin Payments, Cybersecurity Insights, Bitcoinheist, Data Visualization, Power BI, Attack Patterns, Financial Impact.

I. Introduction

Ransomware has become one the most widespread and costliest cybersecurity threats, affecting organizations and individuals around the globe. The rise of cryptocurrency, particularly Bitcoin, has transformed the landscape of these attacks, offering cybercriminals anonymity and the ability to conduct transactions across borders with ease. Among the various forms of ransomware, the Bitcoinheist dataset provides valuable information on attacks that demand Bitcoin payments, capturing details such as ransom amounts, Bitcoin addresses, attack methods, and the sectors being targeted.Use the enter key to start a new paragraph. The appropriate spacing and indent are automatically applied.

This study harnesses the power of Power BI to turn the Bitcoinheist data into a visual story that highlights key metrics like the number of unique attacks, total ransom payments, and the distribution of Bitcoin addresses. By analyzing these metrices, we aim to uncover significant trends and patterns in how ransomware demands over tie. While sector specific insights shed light on which industries are particular vulnerable to these attacks.

The use of interactive visualizations such as geospatial maps, network diagrams, and time based charts offers cybersecurity experts and decision makers a clearer, more detailed view of the ransomware threat. Ultimately, the insights derived from this analysis will help improve strategies for preventing, detecting, and responding to ransomware, equipping organizations and authorities with the knowledge needed to combat these evolving threats more effectively

II RELATED WORKS

The rise of ransomware, especially attacks that demand Bitcoin payments, has been the focus of much research in recent years. Many studies have explored the growing prevalence of these attacks. The financial impact they have, and the challenges of defending against them.

One area of research focuses on the financial implications of ransomware. Reasearchers have highlighted how Bitcoin has become the currency of choice for cybercriminals due to its anonymity, making it harder to trace illicit transactions. Studies by Anderson et al. (2019) show that use of cryptocurrency allows the attackers to evade detection and complicates efforts by law enforcement to stop these crimes. The difficulty to tracking Bitcoin transactions means that it's hard to fully assess the economic damage caused by these attacks.

Another key area of exploration is the attack methods and industries targeted by ransomware. Research by Khan et al. (2018) and Cappelli et al. (2021) had identified certain sectors like Healthcare, Finance, and government that are especially vulnerable to ransomware attacks. These industries often hold sensitive data or provide critical services, making them prime targets. The impact of an attack on these sectors can be devasting, which is why they are frequently targeted by cybercriminals.

The role of data visualization and analytics tool in understanding ransomware trends has also been a focus. Research by Liao et al. (2020) and Rao et al. (2021) shows that using interactive tools like Power BI can provide cybersecurity experts with a clearer view of ransomware activity. These tools allow users to explore the date in various ways such as mapping out of attack hotspots, tracking changes over time, and uncovering connections between Bitcoin addresses involved in different campaigns. Such visualizations can make it easier to understand the scope of the problem and spot patterns that would otherwise be difficult to detect.

III METHODOLOGY

a) Dataset description

This dataset provides a structured representation of unique entities identified by alphanumeric addresses, with associated features describing their characteristics, relationships, and behaviors over time. It is designed for analytical and predictive purposes, particularly in understanding patterns, classifying entities, and exploring connections. The data offers rich insights into temporal, quantitative, and categorical attributes, making it a valuable resource for research in domains such as data science, behavioral analysis, and anomaly detection.

b) Applications and research potential

This dataset is primed for exploring relationships, behaviors, and classifications in a networked system. It supports various research goals:

- Anomaly Detection: Identifying unusual patterns or outliers based on attributes like "Weight," "Neighbors," or "Income."

- Predictive Modeling: Leveraging features and labels to predict outcomes or behaviors using machine learning models

Network Analysis: Examining connectivity patterns and entity influence through metrics like "Looped" and "Neighbors."

- Time-Series Analysis: Uncovering trends or seasonality using "Year" and "Day."

c) Existing systems in ransomware analytics

Ransomware analytics has been an evolving field with diverse approaches developed in the past, leveraging various datasets, tools, and techniques. The existing systems can be broadly categorized into data analysis platforms, blockchain tracing tools, and network-based detection systems:

1. Data Analysis Platforms: Many existing systems rely on statistical tools and data visualization software, such as Power BI, Tableau, and Python-based frameworks, to analyze ransomware datasets. These systems provide insights into attack patterns, ransom demands, and victim profiles by processing historical data. They are useful for visualizing trends but often lack real-time analytics or predictive capabilities.

2. Blockchain Tracing Tools: Tools like Chainalysis and Elliptic are widely used for tracing cryptocurrency transactions. These platforms excel at tracking the flow of Bitcoin payments, identifying suspicious wallets, and linking transactions to potential ransomware campaigns. However, their reliance on publicly available blockchain data can be limited by the anonymity features of cryptocurrencies.

3. Network-Based Detection Systems: Intrusion detection systems (IDS) and intrusion prevention systems (IPS) are designed to monitor network traffic for malicious activities. These systems, powered by machine learning or rule-based algorithms, aim to detect ransomware behavior such as data exfiltration or encryption patterns. However, their effectiveness is often hindered by sophisticated evasion techniques employed by ransomware actors.

d) Drawbacks identified

1. Lack of Real-Time Insights: *Most existing systems focus on retrospective data analysis, which delays the identification of ongoing attacks. Without real-time processing, these systems fail to provide actionable intelligence during live incidents.*

2. Anonymity of Cryptocurrencies: *Blockchain tracing tools face challenges in fully de-anonymizing Bitcoin transactions. Advanced obfuscation techniques, such as mixing services and multi-signature wallets, make it difficult to trace the origins of payments or link addresses to specific ransomware campaigns.*

3. Fragmented Data Sources*: The integration of diverse data sources, such as ransomware samples, victim profiles, and financial transactions, remains limited. This fragmentation prevents a holistic analysis of ransomware campaigns and weakens the predictive potential of these systems. 4. Insufficient Attack Vector Analysis: Existing systems often fail to correlate ransomware entry points, such as phishing or software vulnerabilities, with the scale and impact of attacks. This hinders organizations from prioritizing specific defenses.*

e) Proposed solution

To address the challenges and limitations identified in ransomware analytics, a comprehensive and proactive solution is proposed. This solution leverages advanced analytics, machine learning, and interactive visualizations to enhance the understanding, detection, and mitigation of ransomware activities. By utilizing the Bitcoinheist dataset and integrating with external intelligence sources, the system aims to provide actionable insights for cybersecurity professionals, policymakers, and law enforcement.

The proposed solution involves using Power BI to transform the Bitcoinheist data into interactive and dynamic visualizations that reveal critical trends and connections. Key performance indicators (KPIs) such as the total number of Bitcoin addresses involved, the average ransom amounts, and the total ransom payments will be calculated to measure the scope of ransomware activities. These metrics offer a comprehensive view of the financial impact and operational scale of ransomware campaigns.

In addition, a series of charts and diagrams are proposed to uncover deeper insights:

- Geospatial maps will identify regional hotspots of ransomware attacks, helping stakeholders focus resources on the most vulnerable areas.

- Line and column charts will visualize year-over-year changes in attack frequency and ransom amounts, enabling trend analysis and forecasting.

- Donut charts and treemaps will categorize ransomware families and attack vectors, providing an understanding of dominant threats and entry points

- Scatter plots will analyze the relationship between ransom amounts and the likelihood of payment, offering strategic insights for decision-making during incidents.

The solution also incorporates network diagrams to map connections between Bitcoin addresses involved in ransomware activities. This approach identifies patterns of address reuse, links across campaigns, and potential threat actor groups, enhancing the ability to trace and disrupt ransomware operations.

To overcome current limitations, the solution integrates predictive modeling using machine learning algorithms. These models will analyze historical trends and patterns to forecast potential future ransomware activities, improving preparedness and response. By combining these predictive capabilities with interactive visualizations, the proposed system ensures both high-level overviews and granular details are accessible to stakeholders.

Finally, this solution emphasizes the need for collaboration between organizations, law enforcement, and policymakers. Through secure data-sharing mechanisms and integrated analytics, the system fosters collective efforts to prevent, detect, and mitigate ransomware threats. This holistic approach aims to reduce the financial and operational impact of ransomware, while strengthening defenses against evolving threats.

f) Proposed architecture for power bi ransomware analytics system

1. Data Collection Layer

In this layer, data is gathered from multiple sources to form the dataset used for analysis in Power BI. This involves the collection of raw ransomware data and external threat intelligence to enrich the analysis.

- Bitcoinheist Dataset: The core data containing information about ransomware attacks, Bitcoin addresses, ransom amounts, attack vectors, and victim industries.

- External Data Sources: Data from public threat intelligence feeds, such as blockchain analysis tools (e.g., Chainalysis, Elliptic), law enforcement incident reports, or opensource cybersecurity reports.

- Data Feeds: Any live or periodic data sources providing up-to-date ransomware activities to continuously enrich the dataset.

2. Data Processing Layer

This layer is responsible for preparing the raw data for analysis. Power BI supports ETL (Extract, Transform, Load) operations using its Power Query functionality.

- Data Extraction: Data is extracted from various sources such as Excel files, CSVs, or APIs from blockchain monitoring tools.

- Data Transformation: Using Power Query in Power BI, raw data is cleaned, normalized, and structured. This step ensures the data is in a format suitable for visualization and reporting.

- Data Aggregation: The data is aggregated at the level required for reporting— calculating

3. Data Modeling Layer

In this layer, Power BI's data model is created. It involves defining tables, relationships, and key metrics (KPIs) for analysis.

- Data Modeling in Power BI: Fact Tables: These contain transactional data like ransom payments, attack incidents, or Bitcoin addresses involved in attacks. Dimension Tables: These categorize data (e.g., ransomware families, geographical regions, attack vectors, and sectors). Calculated Columns: New metrics like total ransom amount, average ransom amount, and attack frequency are calculated.

4. Analytical Layer

Here, the analysis is carried out using Power BI's visualizations and DAX (Data Analysis Expressions) for detailed reporting.

- Visualizations: A variety of visual elements are used to explore ransomware trends. These include:

KPI Cards: Display total Bitcoin addresses involved, total ransom amounts, and average ransom values.

Bar and Line Charts: Used to track trends over time, such as ransom demands or attack frequency.

Geospatial Heatmaps: Geographical distribution of attacks.

Treemaps and Donut Charts: Used to visualize ransomware families and attack vectors

Network Diagrams (using Power BI's custom visuals): Visualize connections between Bitcoin addresses involved in multiple campaigns.

5. User Interface Layer

This layer provides the interface through which users interact with the system.

- Power BI Dashboard: Interactive and user-friendly dashboards where users can drill down into specific ransomware attacks, explore metrics by time, region, or attack vector, and analyze trends over time.

- Filters and Slicers: Users can filter by ransomware family, sector, region, attack type, or time period for detailed exploration.

- Real-time Data: The system can be designed to refresh periodically to keep the visualizations up-to-date with new data.

g)Algorithms and techniques used in power bi

While this solution does not involve machine learning algorithms, complex graph algorithms, or other predictive models, several techniques and methods are used within Power BI to drive analysis and insights:

1. DAX (Data Analysis Expressions)

DAX is a powerful formula language used in Power BI for creating calculated columns, measures, and aggregations. Some examples of DAX calculations used in this project:

- Total Ransom Amount: A simple DAX formula that sums up all the ransom payments across incidents.

- Average Ransom Amount: Calculated by dividing the total ransom amount by the number of incidents.

2. Time-Series Analysis

Power BI allows users to create time-series visualizations, and DAX formulas can be used to analyze changes in attack frequency and ransom amounts over time. This enables users to track patterns, identify peak periods, and understand seasonal trends in ransomware attacks.

3. Geospatial Analysis

Power BI provides built-in mapping and geospatial visualization tools. Through the use of custom visuals (such as ArcGIS Maps for Power BI), the system can:

- Plot ransomware incidents on a map, highlighting regional hotspots.

- Use geospatial clustering to identify areas with high attack frequency.

4. Data Aggregation

Data aggregation in Power BI helps summarize detailed data into meaningful insights. For example, aggregation of ransom amounts and attack types helps identify which sectors are most affected and which types of attacks are more prevalent. Aggregations are done using DAX formulas like SUMX, AVERAGEX, or COUNTROWS.

5. Network Diagrams (Power BI Custom Visuals)

Power BI supports custom visuals, such as Network Navigator or Force-Directed Graphs, which can be used to visualize relationships between Bitcoin addresses. These diagrams map connections between addresses involved in multiple ransomware attacks, helping to trace the flow of ransom payments.

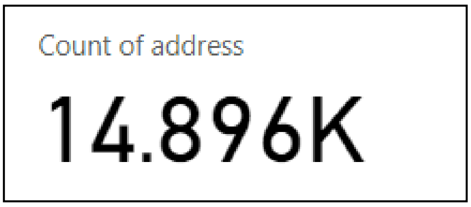
6. Interactive Dashboards and Filters

The use of filters and slicers enables the creation of interactive dashboards, where users can dynamically explore the data. For example:

- Users can filter by ransomware family to view which families are the most active.

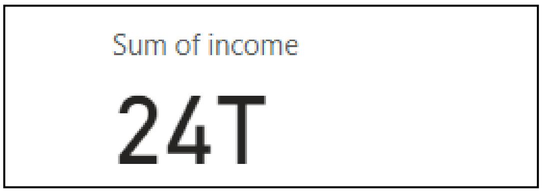
- Filters can also be applied to view specific time frames, such as yearly trends or attack trends by region.

IV RESULTS AND DISCUSSIONS


Count of address

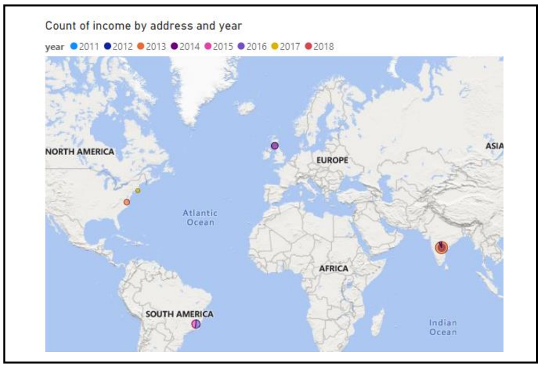
14.896K

14.896K bitcoin addresses shows the unique addresses linked to ransomware payments, indicating the scale of the ransomware campaign. Compared to existing models, this approach offers real-time insights, scalability, and granular analysis, enabling more effective tracking and monitoring of ransomware activities.

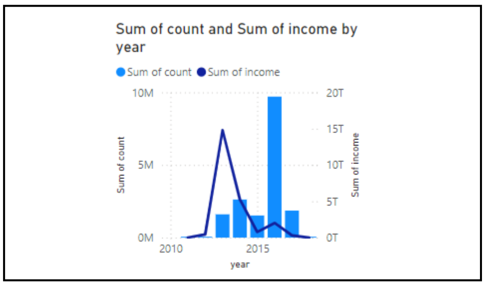

Sum of income

24T

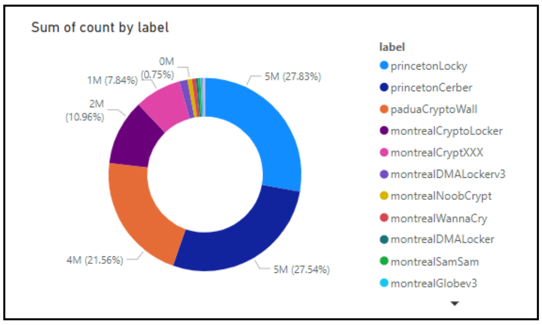
Total ransom amount: sum of all recorded ransom payments in Bitcoin


Count of income by address and year
year ● 2011 ● 2012 ● 2013 ● 2014 ● 2015 ● 2016 ● 2017 ● 2018
NORTH AMERICA
EUROPE
ASIA
Atlantic Ocean
AFRICA
SOUTH AMERICA
Indian Ocean

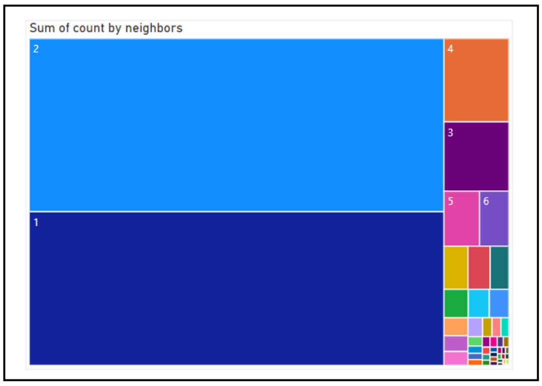
This map chart visualizes the count of income (likely ransom payments) by address and year, spanning from 2011 to 2018. The data points are color-coded by year, allowing for easy identification of ransomware activity over time and regions.



line and clustered column chart to visualize year-on-year changes in ransomware attack frequency and ransom amounts



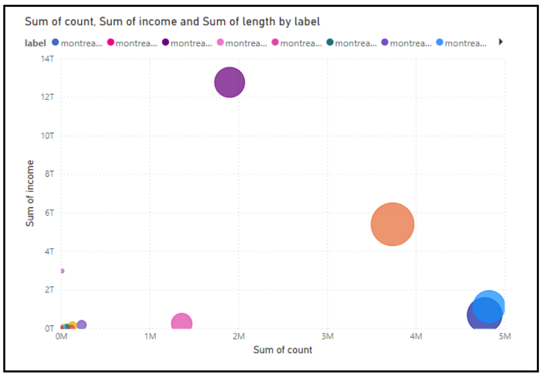
a donut chart to visualize the percentage of attacks by ransomware family or variant.



A TREE MAP CHART TO VISUALIZE RANSOMWARE DISTRIBUTION BY ATTACK VECTOR (E.G., PHISHING, RDP EXPLOITATION, SOFTWARE VULNERABILITIES).

Sum of income by address and neighbors

network diagram to visualize connections between bitcoin addresses associated with different ransomware campaigns.



a bubble chart to visualize ransomware impact factors by number of affected systems, ransom amount, and duration of system downtime

CONCLUSION

This paper presents an innovative approach to tracking and analyzing ransomware activities through Power BI, focusing on Bitcoin addresses involved in ransom payments leveraging Power BI's powerful data modeling and visualization capabilities, we provide real-time insights, scalability, and an interactive platform for identifying trends and patterns in ransomware campaigns. The system offers a comprehensive, user-friendly interface that enhances the understanding of ransomware distribution, its financial impact, and the network of addresses involved. It highlights the importance of data-driven decision-making in combating ransomware threats effectively.

a)Future scope

Machine Learning Integration: Future versions could integrate machine learning algorithms to predict ransomware attack patterns based on historical data and blockchain analysis, enhancing proactive threat detection.

- Real-time Threat Detection: The system could evolve to provide real-time alerts, identifying new ransomware addresses as soon as transactions occur.
- Enhanced Geospatial Mapping: Incorporating advanced geospatial features could provide deeper insights into regional ransomware trends and help law enforcement agencies track and respond faster.
- Broader Data Sources: The system could include additional data sources, such as dark web monitoring, to offer a more comprehensive picture of ransomware activities.
- Blockchain Analytics: Incorporating advanced blockchain analysis tools to track ransom payments and identify suspicious activity could further enhance the system’s capabilities

b) References

Blockchain analysis tools like Chainalysis and Elliptic for tracking ransomwarerelated transactions.

- Power BI Documentation for advanced visualizations and DAX functions used in data analysis and reporting.
- Past research on ransomware trends and Bitcoin usage in cybercrimes for foundational understanding and methodology.
- Existing cybersecurity reports from organizations such as Symantec and Kaspersky regarding ransomware campaigns and Bitcoin use in cybercrime.