

# Graph Visualization of Cyber Threat Intelligence Data for Analysis of Cyber Attacks

Mücahit Sülü and Resul Daş


**Abstract**—Threat intelligence enables us to make faster, more informed, data-backed security decisions and change their behavior from reactive to proactive in the fight against threat actors. Cyber threat intelligence sources include open-source intelligence, social media intelligence, human intelligence, technical intelligence, device log files, forensically acquired data or intelligence from the internet traffic, and data derived for the deep and dark web. In this study, graph visualization is discussed for the intelligible and accurate analysis of complex cyber threat intelligence data, including network attacks. The processes of collecting, cleaning, organizing, and visualizing cyber intelligence data in different formats and contents on a single platform are given step by step. Dynamic graphs play an active role in these systems, where the attack locations and targets from different points are constantly variable. Therefore, research on dynamic graph solutions and visualization in the visual analysis of cyberattacks is presented.


**Index Terms**—cyber security, graph visualization, dynamic graph, cyber threat intelligence, cyber attack visualization, big data.

## I. INTRODUCTION

GRAPHS HAVE created models for solving and analyzing many unsolved problems throughout history. These models have a wide variety of types and formats. Graph visualization is of great importance as it provides convenience in terms of understanding and tracking problems. Today, graphs have become so popular that it is possible to encounter a problem or an algorithm modeled with graphs in almost every field in the scientific world. Effective solutions through graphs are seen in the literature in almost all fields such as computer sciences [1], social sciences[2], linguistics [3], engineering[4], mathematics [5], biology and genetics [6].

Every day in meeting rooms, people use graph techniques to label relationships and create diagrams to explain their thoughts to others. Graphs can express relatively complex concepts that other visualizations cannot. Data expressed with graphs can be analyzed with graph techniques. The right technique, when chosen wisely, can give the simplest and most intuitive expression of a particular type of knowledge. When poorly selected, a graph can be painfully abstract and broad. Graph visualization may seem confusing. The standard

 **Mücahit SÜLÜ** is with the Department of Computer Programming, Organized Industrial Zone Vocational School, İnönü University, Malatya, 44280 TÜRKİYE e-mail: [mucahit@inonu.edu.tr](mailto:mucahit@inonu.edu.tr)

 **RESUL DAŞ** is with the Department of Software Engineering, Technology Faculty, Firat University, Elazığ, 23119 TÜRKİYE e-mail: [rdas@firat.edu.tr](mailto:rdas@firat.edu.tr)

Manuscript received March 19, 2022; accepted July 28, 2022.  
DOI: [10.17694/bajece.1090145](https://doi.org/10.17694/bajece.1090145)

pie charts and bar charts you would normally find in a cyber control panel are not the subject of graph theory. Graph theory is the visualization of connections and relationships in data. Cyber attacks can be made to any device connected to a local network or internet. It is vitally important to detect these attacks and take action accordingly. Graph visualization methods and graph techniques can be used to make sense of and analyze cyber attacks.

In this article, the steps of analyzing cyber threat intelligence data in different formats through various stages and visualizing them as graphs are examined. It is a concise study that guides researchers working in this field. In this context, the second part of the article examines the cyber intelligence data. In the third part, graph visualization process steps and processes are explained. In the last part, the general results of the study are presented.

## II. BACKGROUND

Intelligence is tactical, technical or predictable information that is specially collected and analyzed through certain filters for presentation to military or political higher authorities. We can describe it as valuable data that has been unearthed by combining interrelated parts over raw information obtained from almost any source. Intelligence activities are indispensable for states and require the processing of information and documents compiled from various sources in response to the needs determined by the state. In history, the effectiveness of intelligence data has always come to the fore in the destruction or establishment of states, and in winning or losing wars. For this reason, almost all heads of state have made an effort to collect sound and reliable intelligence data and use them in the most effective way. When we take a look at Turkish history, we see that Sultan Abdülhamit II was a sultan who understood the necessity of intelligence and applied it, and benefited from intelligence activities by establishing the Yıldız Intelligence Organization. Additionally, we see in history that he took a cautious approach towards the intelligence data and thus, made decisions after verifying the reliability of the source and the integrity of the data. Considering the types of intelligence according to their fields, there are various types of intelligence data such as military, political, economic, geographical, social, biographical, technological, transportation and communication. However, in this study, we focus on the content, scope, analysis and visualization of cyber intelligence data within the scope of technological intelligence data.

### A. Cyber Intelligence

Cyber intelligence is the collection and discovery of threats from electronic media that can harm institutions and organizations, business elements and security at any level. It is a type of intelligence that enables early measures to be taken by detecting the aims, methods or attack types of the attackers as a result of analyzing the data collected and enriched from electronic media through a process. When we look at the data breaches we have encountered in recent years, it has been revealed that the measures taken during or after the cyber attacks do not always work. Due to the cyber world we live in and the rapid development of technology, institutions can face different threats and can receive hundreds of attacks at any time. It is not easy to follow the cybercriminals and their techniques targeting critical systems in the internet world, and it is a situation that requires large budgets and dealing with big data. At this point, cyber intelligence emerged and started to play an important role in cyber attacks. While cyber attackers carry out their cyber activities in a highly motivated way, they improve their attack methods and diversity day by day. Because of this increasing cyber threat, it is critical to be aware of an attack before it happens. The power of predicting attacks accelerates the decision-making process of institutions. At that point, the importance of Cyber Intelligence is increasing day by day.

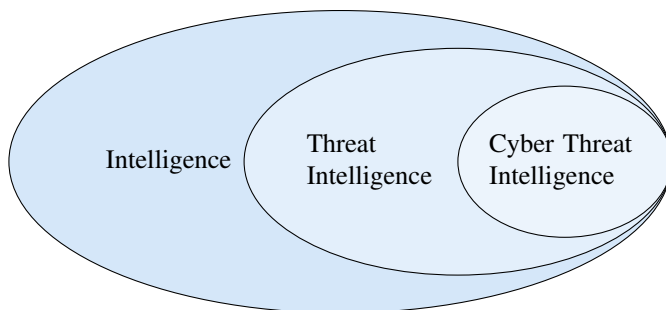


Fig. 1. Cyber Threat Intelligence

### B. Cyber Threat

A cyber threat is an attempt by malicious persons or organizations to gain unauthorized access to control system devices or network, disrupting the network structure or rendering it unusable. Cyber threats can originate from various different points/places, people, institutions or organizations. This is where cyber intelligence comes into play. The actions taken by cyber threat sources such as terrorists, hackers, commercial competitors, spies, hostile states, unhappy employees, organized crime groups with the aim of causing harm are called cyber threats. These threats provide insight into what kind of scenario attackers might follow when attacking their victims. Malware, Spyware, Malvertising, Man in the Middle (MITM), Wiper Attacks, Distributed Denial of Service (DDoS), Ransomware, Botnets, Trojans, Phishing, Data Breaches, Worms, Keyloggers, Backdoors, Advanced Persistent Threats are important examples of cyber threats.

### C. Cyber Threat Intelligence

Cyber threat intelligence - CTI) is knowledge, skills, and experience-based information concerning the occurrence and assessment of both cyber and physical threats and threat actors that are intended to help mitigate potential attacks and harmful events occurring in cyberspace. The purpose of cyber threat intelligence is to help institutions and organizations understand the risks of cyber attacks or cyber threats. Examples of these attacks are 0-day attacks, crypto viruses, APT (Advanced Persistent Threat), botnets or exploits. These threat elements are reported with the intelligence activities revealed using various software tools and presented to the relevant institutions and organizations together with special protection methods, thus providing guidance for an active defense. Such attacks can cause serious damage to institutions and organizations. Thanks to cyber threat intelligence, extensive and deep analysis data is used to help protect institutions and organizations from such attacks.

1) *Cyber Threat Intelligence Data*: Advances in attack methods make it extremely difficult to identify the attacker and the attack. Traditional security measures such as firewalls, signature registration, and intrusion detection system (IDS) fail to prevent these new types of attacks. To meet these challenges, the emerging field of cyber threat intelligence uses artificial intelligence and machine learning techniques to intelligently detect, learn and overcome advanced cyber attacks. There is an increasing trend in the use of Machine Learning (ML) and data mining techniques in the static and dynamic analysis of malware, due to their efficiency and powerful network anomaly detection. In addition to these, different mechanisms such as honeypot are used to deceive the attackers. In such methods, security professionals use fake information or sources that appear to be legitimate to lure attackers, monitor the attackers' activities, and detect the attack and its type. Cyber threat intelligence data mining has an increasing popularity today.[7]

TABLE I  
CYBER OBSERVABLE SPECIAL DATA TYPES [8]

Type	Description
boolean	True or False.
float	One IEEE 754 [ IEEE 754-2008] double-precision number.
hashes	One or more cryptographic hashes.
integer	Integer.
list	An ordered array of values.
open-vocap	Type from a STIX or suggested word value.
string	Unicode character string.
timestamp	A time value (date and time).
binary	A byte array.
hex	A decimal number at the base of eight.
dictionary	Set of key-value pairs.
object-ref	Cyber observable reference.
observable-objects	One or more cyber observable objects.

The data is collected, analyzed and organized from the deep/dark web, blogs, social media and forums, often with artificial intelligence support. Although there are many companies that provide this data today, companies that attach importance to security provide it themselves. Structured Threat Information Expression (STIX) is an application that enables organizations to share cyber threat intelligence data with each other in a consistent and readable way. An example of STIX data is given by the list 1. Cyber observable custom data types are given in Table I. The Figure 7 shows STIX domain object relationships. Here, many relationship objects such as targets, users, viruses, threats are visualized. Figure 8 shows a simple cyber attack graph structure.

```

1 {"type": "bundle",
2  "id": "bundle--5d0092c5-5f74-4287-9642-33f4c354e56d"
3  ,
4  "spec_version": "2.0",
5  "objects": [{
6    "type": "indicator",
7    "id": "indicator--8e2e2d2b-17d4-4cbf-938f...",
8    "created_by_ref": "identity--f431f809-377b...",
9    "created": "2021-04-29T14:09:00.000Z",
10   "modified": "2021-04-29T14:09:00.000Z",
11   "object_marking_refs": [{"marking-definition..."}],
12   "name": "Poison Ivy Malware",
13   "description": "This file is part of P",
14   "pattern": "[file:hashes.'SHA-256' =
15   'aec070645fe53ee3b3763059376134f058cc3372...' ]"
16   },{
17   "type": "marking-definition",
18   "id": "marking-definition--34098fce-860f-48...",
19   "created": "2021-09-01T00:00:00.000Z",
20   "definition_type": "tlp",
21   "definition": {
22     "tlp": "green"}
23   }
24 ]}
25 \caption{}

```

Listing 1. A sample snippet from the STIX 2 package.

#### D. Data Processing

There are two types of data collection methods, the first is collected over a period of time and the second is constantly flowing from one device to the next.

1) *Batch Data Processing*: Batch Data Processing is an efficient way to process large amounts of data collected over a period of time. It also helps reduce the operational costs that businesses can spend on their workforce, as it doesn't need dedicated data entry officers to support its operation. It can be used offline and gives administrators full control over when they start processing, whether overnight, on a weekend, or at the end of a pay period.

As with anything, batch processing has a few downsides. One of the biggest problems businesses see is that these systems can be difficult to debug. If you don't have a dedicated IT team or specialist, trying to repair the system when an error occurs can be detrimental and require an outside consultant. In addition, if the business needs quick returns, for example, if the customer searches for a product on an e-commerce website, the customer will have bought that product because the business's system will analyze and offer this customer weeks or days later.

2) *Streaming Data Processing*: Streaming (Real-Time) Data Processing is the process of analyzing data streaming from one device to another almost instantly. This method of continuous calculation takes place as data flows through the system with no mandatory time limitation on output. With nearly instantaneous streaming, systems do not require large amounts of data to be stored. Stream processing is very useful if the events you want to watch happen frequently and are close together in time. It is also best to use if the incident needs to be detected immediately and responded to quickly. It is also useful for tasks such as stream processing, fraud detection, and cybersecurity. If transaction data is stream processed, fraudulent transactions can be identified and stopped before they are completed. In addition, instant suggestions can be offered to the customer. One of the biggest challenges organizations face with stream processing is that the long-term data throughput rate of the system must be the same or faster than the long-term data-input rate, otherwise the system will start to experience storage and memory related problems. Another challenge is trying to find the best way to deal with the huge amount of data being generated and moved. To keep the data flow in the system at an optimal level, organizations need to create a plan for how to reduce the number of copies, how to target the compute cores, and how to make the best possible use of the cache hierarchy.

### III. DYNAMIC COMPLEX GRAPHS AND SOME TYPES

It is difficult to represent a time dimension with graphs. It is accepted in scientific circles that simultaneous representations of states are more informative than sequential representations. For example, a time series bar chart is much better at showing behavior over time than animating changes in a single bar. In the latter case, repeatedly scrolling the animation back and forth in time would be a means of getting the essence of the change; however, it will not be easy to detect correlations in behavior when comparing, as it will not be as accessible as a time series. It is more effective to be able to see the values over time simultaneously. However, it is not clear how this principle can be applied to graphs. Since data in dynamic graphs changes over time, visualization can be made with data collected by batch data processing (given by II-D1) or Streaming (Real-Time) Data Processing (given by II-D2). Now we will examine their advantages and disadvantages.

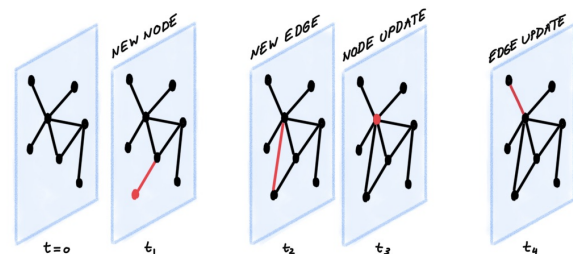


Fig. 2. Dynamic graph sample [9]

**A. Implementation of Dynamic Graphs**

A dynamic graph evolves and changes over time, so it can be viewed as a timed sequence of events. Figure 2 shows an example of a dynamic graph that changes over time. In this example, a new node and a new edge are added at  $t_1$  moment, only a new edge is added at  $t_2$  moment, the node degree is updated at  $t_3$  moment, and the edge is updated at  $t_4$  moment. In the edge update, the weight assigned to the edge can be changed, a label can be assigned or a new edge can be added. In a node update, it can be a node's degree or a label update.

$V$  (set of nodes),  $E$  (set of edges),  $f$  (weights of nodes) and  $g$  (weights of edges) [10];

- In a node dynamic graph, the set  $V$  changes with time. Therefore, some nodes can be added or removed. When the nodes are removed, the edges formed with them are also removed.
- In an edge dynamic graph, the set  $E$  changes with time. Thus, edges can be added or removed from the graph.
- In a node-weighted dynamic graph, the  $f$  function changes with time; so the weights at the nodes also change.
- In an edge-weighted dynamic graph, the  $g$  function changes over time.
- In a full-weighted dynamic graph, both  $f$  and  $g$  functions may change over time.

that are drawn in thinner indicate that the migrations are less intense.



Fig. 4. An example of circle graph. [14]

With Figure 5, the hierarchy of task dependencies in Cyber entities is given, the grouping operation here makes the graph more understandable.

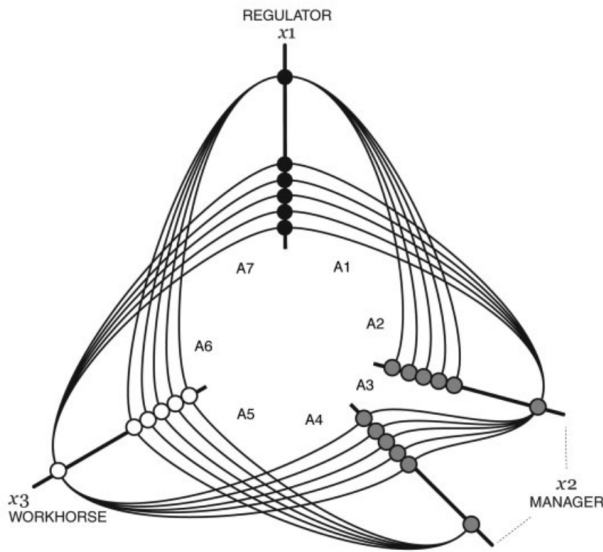


Fig. 3. An example of Network Hive Graph [11]

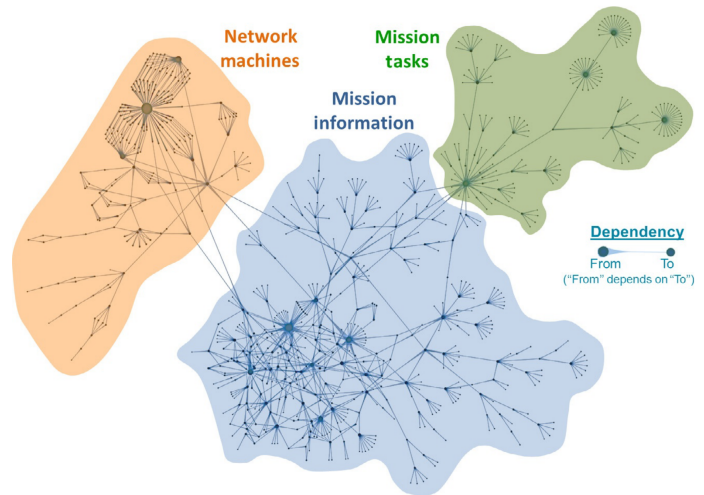


Fig. 5. Hierarchy of mission dependencies on cyber assets. [13]

**B. Types of Some Complex Graphs**

Graphs are a great way to visualize data and display statistics. Popular graph types include Hive Graphs[11], Circle Graphs[12], Hierarchy Graphs [13]. For illustrate, Hive graphs transparently models the network structure, are easy to understand and can be easily modified to identify patterns of interest. An example network structure model is given in Figure 3. In Figure 4, estimates of migration flows between and within regions for the period from 2005 to 2010 are given as a circle graph. In the graph given in Figure 4, the lines drawn in bold indicate that the migrations are more intense, and the lines

**IV. GRAPH VISUALIZATION PROCESS STEPS OF CYBER THREAT INTELLIGENCE DATA**

In order for cyber intelligence data to be analyzed effectively, it must first be expressed as graphs. There are two major challenges for viewing and analyzing potentially very large and complex graphs of cyber attacks. The first is the real-time processing of large amounts of data and converting it into visual format. The second is to visualize complex graphs, including all possible attack paths, by keeping them manageable.[15] In order to visualize complex graphs,

graph visualization steps given in figure 6 are applied and data is made meaningful. These data can then be analyzed using graph analysis methods. Graph algorithms[16] and graph partitioning[17], [18] are some of the graph analysis methods.

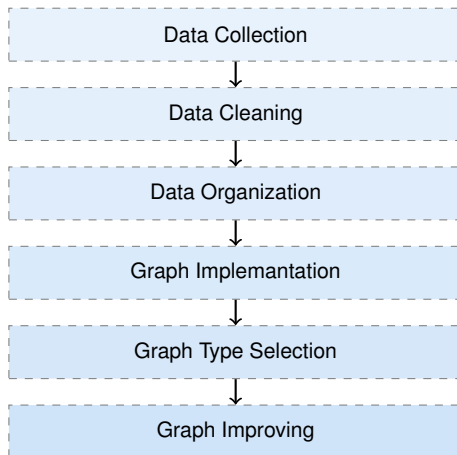


Fig. 6. Steps of graph visualization of cyber threat intelligence data

#### A. Data Collection

Threat intelligence is data that is collected, processed, and analyzed to understand a threat actor's motives, targets, and attack behaviors. Obtaining large-scale cyber intelligence data in real time is a challenging process. Especially in a network system, it is more difficult to detect the attack data from the data traffic flowing and to capture and collect only the relevant parts. In addition, the intelligibility of these data is another challenge. Therefore, in the collection of cyber intelligence data, it is first necessary to determine what type of data will be collected from which source or sources. For this, it is necessary to have a deep and comprehensive networking knowledge. One of the frequently repeated quotes in network analysis is this: Before you connect the dots, you need to collect them. Your first challenge may be to determine what data to collect.

Cyber threat intelligence data can be versatile according to the threat element. These datasets usually contain such as URL, host, IP address, e-mail account, hashes (MD5, SHA1, and SHA256), common vulnerabilities and exposures, registry, file names ending with specific extensions, and the program database path. This information can be found in different log files kept by the server and network devices [19].

#### B. Data Cleaning

After collecting complex and large-scale cyber threat intelligence data, the next step is to clean the data and make it usable. Cleaning up text-based cyber intelligence data in different formats, large sizes and containing various parameters is a very difficult process. This data, where log files are actively used, is frustratingly messy. Unfortunately, most graph software tools are not designed to work with this messy data and must be cleaned and prepared before reading the data into graph software.

The following points should be provided regarding the elimination and cleaning of the difficulties experienced in the cleaning process of the complex data obtained;

- *Inconsistent Node Names*: A node should not be represented by more than one name.
- *Refreshed Nodes*: Each node should appear only once in the node dataset.
- *Refreshed Edges*: Some types of graph visualization and analysis software do not work well with multiple edges between the same pair of nodes and need to be consolidated.
- *Self-Loop*: Some graph software do not handle self-loops.
- *Isolated Node*: Datasets may have nodes with no connections, disconnected nodes may cause problems with graph visualization.
- *Edges Connected to Non-Existing Nodes*: In some datasets, an edge can be defined between two nodes where one of the nodes is not in the node list.
- *Invalid Data*: Real world data can have null or invalid data. The numeric data column can have text entries such as N/A or ERROR. These entries should be cleaned or removed.
- *Units*: All numeric data needs to be normalized to the same numeric units.

There are many approaches to correcting invalid, incomplete and inconsistent data. A simple approach would be to remove the specific problematic record, but it may be more beneficial to use other approaches that involve entering missing values or normalizing the data. Cyber attack data is generally not regular data because it is very repetitive and different methods (VPN, etc.) are used to deflect the source. In addition to the steps we mentioned above to clean this data, the relationships must be captured correctly so that the source device or devices can be expressed clearly.

#### C. Data Organization

It is very effective to define and organize complex and big data into a set of nodes and edges. This clear separation will enable data exploration with a wider variety of tools. Cyber-attack data consists of attacking devices, attacked devices and their relationships and attacks. For example, we can organize our data by expressing the attacking and the attacked devices as nodes, the relationships between these devices and the attack data as weighted edges.

#### D. Graph Implementation

After the graph data is cleaned and organized, it can be analyzed with software tools [21], [22], [23] where graph analyzes are made, or it can be analyzed and visualized by coding with programming languages. The next goal is to make the graph more understandable. In this context, the following questions are examined and the process proceeds.

- Are the assigned nodes all interconnected or in many separate pieces?
- Is the resulting graph a hierarchy?
- Is the resulting graph sparse? Or is it heavily linked?
- Are there any obvious clusters in the resulting graph?

Statistics can provide large and multidimensional information using data and answer questions about size, density, and

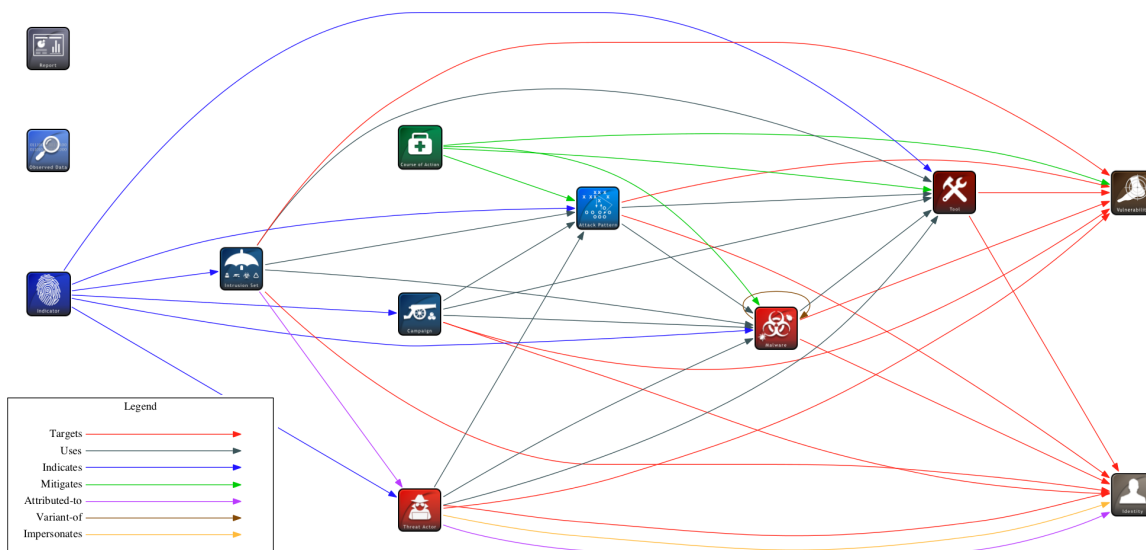


Fig. 7. Visualized STIX domain object relationships [20]

number of discrete graphs. Layouts are an important visual technique for understanding graph structure. Different layouts will reveal different aspects of the graph, allow for different types of analysis, and support different types of stories. A wide variety of node and link layouts can provide different ways of revealing links, groupings, and sequences in graphs. Other graph layout types focus on other properties of the graph, displaying lowest values, hierarchies, or multiple attributes.

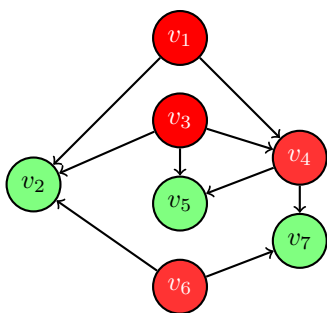


Fig. 8. A simple cyber attack graph

*E. Graph Type Selection*

During the graph type selection phase, graph images suitable for the type, property and size of the data can be determined. Graph types are increasing day by day. These can be graph visualization types such as Circle Graph [12], Sankey Diagram [24], Hive Graph[11], it is necessary to choose the graph that is most suitable for our data. We can choose different types of graphs according to the desired features in highlighting the data. For example, when cyber threat intelligence data is expressed as a hive graph, only attack density can be analyzed. Relationships can be seen more clearly when using the circle graph.

*F. Graph Improving*

In the graph visualization process, since we have chosen a graph that is suitable for our data, we can now improve the features we want to emphasize, with methods such as coloring, ghost effect, fading, labeling so that our graph can become more understandable. In this processes, the symbols of the related device or devices can be used in the symbols of the nodes. Thus, the graphs to be created will be more meaningful and more understandable by people who are not closely interested in the subject.

V. CONCLUSION

Graphs are of great importance as they form a model for solving and analyzing many unsolved problems in history. Today, graphs have become so popular that it is possible to encounter a problem or an algorithm modeled with graphs in almost every field. We encounter graphs in almost all fields such as computer science, social sciences, linguistics, engineering, mathematics, and medicine. Cyber threat intelligence benefits organizations of all shapes and sizes by helping process threat data to better understand their attackers, respond faster to incidents and proactively get ahead of a threat actor’s next move. For SMBs, this data helps them achieve a level of protection that would otherwise be out of reach. On the other hand, enterprises with large security teams can reduce the cost and required skills by leveraging external threat intel and making their analysts more effective. In this study, the visualization of complex cyber threat intelligence data as graphs is examined for accurate analysis of network attacks. Within the scope of graph visualization, colouring, segmentation and presenting with different patterns is a very important and technical issue. In this study, the graph visualization processes of cyber threat intelligence data, in which cyber attacks can be detected, are presented in detail. This study opens the floodgates to the monitoring and visualization of network attacks in the Internet environment, the world’s largest information network. Further

studies could be done on the acquisition of critical real-time data and its visualization with dynamic graphs.

## REFERENCES

- [1] Y. Bürhan and R. Daş, "Co-author link prediction from academic databases," *Gazi University Journal of Polytechnic*, vol. 20, no. 4, pp. 787–800, Dec. 2017. [Online]. Available: <http://dergipark.gov.tr/download/article-file/387477>
- [2] L. Yang, E. Cheng, and Z. M. Özsoyoğlu, "Efficient path-based computations on pedigree graphs with compact encodings," *BMC Bioinformatics*, vol. 13, no. S3, p. S14, Dec. 2012. [Online]. Available: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-13-S3-S14>
- [3] Q. Guo, X. Qiu, X. Xue, and Z. Zhang, "Syntax-guided text generation via graph neural network," *Science China Information Sciences*, vol. 64, no. 5, p. 152102, May 2021. [Online]. Available: <http://link.springer.com/10.1007/s11432-019-2740-1>
- [4] B. Xie, C. Qi, H. Ben, and W. Yu, "The applications of graph theory in electric network," in *2019 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC)*. Beijing, China: IEEE, Aug. 2019, pp. 780–784. [Online]. Available: <https://ieeexplore.ieee.org/document/9168962/>
- [5] D. P. Sinha, "A pairing between graphs and trees," *arXiv:math/0502547*, Oct. 2006, arXiv: math/0502547. [Online]. Available: <http://arxiv.org/abs/math/0502547>
- [6] S. A. M. A. Junid, N. M. Tahir, Z. A. Majid, and M. F. M. Idros, "Potential of graph theory algorithm approach for DNA sequence alignment and comparison," in *2012 Third International Conference on Intelligent Systems Modelling and Simulation*. Kota Kinabalu, Malaysia: IEEE, Feb. 2012, pp. 187–190. [Online]. Available: <http://ieeexplore.ieee.org/document/6169697/>
- [7] J. Zhao, Q. Yan, J. Li, M. Shao, Z. He, and B. Li, "TIMiner: Automatically extracting and analyzing categorized cyber threat intelligence from social data," *Computers & Security*, vol. 95, p. 101867, Aug. 2020. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0167404820301395>
- [8] OASIS, "STIX™ version 2.0. part 3: Cyber observable core concepts."
- [9] E. Rossi, B. Chamberlain, F. Frasca, D. Eynard, F. Monti, and M. Bronstein, "Temporal graph networks for deep learning on dynamic graphs," *arXiv:2006.10637 [cs, stat]*, Oct. 2020, arXiv: 2006.10637. [Online]. Available: <http://arxiv.org/abs/2006.10637>
- [10] F. Harary and G. Gupta, "Dynamic graph models," *Mathematical and Computer Modelling*, vol. 25, no. 7, pp. 79–87, Apr. 1997. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0895717797000502>
- [11] M. Krzywinski, I. Birol, S. J. Jones, and M. A. Marra, "Hive plots—rational approach to visualizing networks," *Briefings in Bioinformatics*, vol. 13, no. 5, pp. 627–644, Sep. 2012. [Online]. Available: <https://academic.oup.com/bib/article-lookup/doi/10.1093/bib/bbr069>
- [12] R. Das and I. Turkoglu, "Creating meaningful data from web logs for improving the impressiveness of a website by using path analysis method," *Expert Systems with Applications*, vol. 36, no. 3, Part 2, pp. 6635–6644, Apr. 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0957417408005952>
- [13] S. Noel, E. Harley, K. Tam, M. Limiero, and M. Share, "Chapter 4 - cygraph: Graph-based analytics and visualization for cybersecurity," in *Cognitive Computing: Theory and Applications*, ser. Handbook of Statistics, V. N. Gudivada, V. V. Raghavan, V. Govindaraju, and C. Rao, Eds. Elsevier, 2016, vol. 35, pp. 117–167. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0169716116300426>
- [14] "Global international migration flows | Wittgenstein Centre." [Online]. Available: [http://download.gsb.bund.de/BIB/global\\_flow/](http://download.gsb.bund.de/BIB/global_flow/)
- [15] G. Chen, "Information fusion and visualization of cyber-attack graphs," *SPIE Newsroom*, 2007. [Online]. Available: <http://www.spie.org/x14562.xml>
- [16] M. Alshammari and A. Rezgui, "An all pairs shortest path algorithm for dynamic graphs," *International Journal of Mathematics and Computer Science*, p. 20, 2020.
- [17] J. R. Nascimento, U. S. Souza, and J. L. Szwarcfiter, "Partitioning a graph into complementary subgraphs," *Graphs and Combinatorics*, vol. 37, no. 4, pp. 1311–1331, Jul. 2021. [Online]. Available: <https://link.springer.com/10.1007/s00373-021-02319-4>
- [18] S. V. Patil and D. B. Kulkarni, "K-way spectral graph partitioning for load balancing in parallel computing," *Bharati Vidyapeeth's Institute of Computer Applications and Management*, Aug. 2021. [Online]. Available: <https://link.springer.com/10.1007/s41870-021-00777-w>
- [19] M. Baykara, R. Daş, and G. Tuna, "Web sunucu erişim kütüklerinden web ataklarının tespitine yönelik web tabanlı log analiz platformu," *Firat Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 28, pp. 291 – 302, 2016.
- [20] "Visualized SDO relationships," Sep. 2021. [Online]. Available: <https://oasis-open.github.io/cti-documentation/examples/visualized-sdo-relationships>
- [21] S. Majeed, M. Uzair, U. Qamar, and A. Farooq, "Social Network Analysis Visualization Tools: A Comparative Review," in *2020 IEEE 23rd International Multitopic Conference (INMIC)*. Bahawalpur, Pakistan: IEEE, Nov. 2020, pp. 1–6. [Online]. Available: <https://ieeexplore.ieee.org/document/9318162/>
- [22] S. Hussain, L. Muhammad, and A. Yakubu, "Mining social media and DBpedia data using Gephi and R," *Journal of Applied Computer Science & Mathematics*, vol. 12, no. 1, pp. 14–20, 2018. [Online]. Available: [http://www.jacsm.ro/view/?pid=25\\_2](http://www.jacsm.ro/view/?pid=25_2)
- [23] G. Drakopoulos, A. Baroutiadi, and V. Megalooikonomou, "Higher order graph centrality measures for Neo4j," in *2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA)*, Jul. 2015, pp. 1–6.
- [24] E. Curmi, R. Fenner, K. Richards, J. M. Allwood, B. Bajželj, and G. M. Kopec, "Visualising a stochastic model of californian water resources using sankey diagrams," *Water Resources Management*, vol. 27, no. 8, pp. 3035–3050, Jun. 2013. [Online]. Available: <http://link.springer.com/10.1007/s11269-013-0331-2>



**Mücahit Sülü** is a lecturer at İnönü University Organized Industrial Zone Vocational School Computer Programming Department. He graduated from the Mathematics Department of the same university in 2007. He continues his doctoral studies at the Department of Software Engineering at Firat University. Current research interests include graph theory, software design and architecture, IoT/M2M applications, software quality and assurance, big data analysis and visualization. He has been actively developing software (web-based automation systems) at İnönü University since 2008.



**Resul Das** is a full professor and Chair in the Department of Software Engineering, Technology Faculty, Firat University, where he has been a faculty member since 2011. He graduated with B.Sc. and M.Sc. degrees from the Department of Computer Science at the Firat University in 1999 and 2002 respectively. Then he completed his Ph.D degree at the Department of Electrical-Electronics Engineering at the same university in 2008. He served as both lecturer and network administrator at the Department of Informatics in Firat University from 2000 to 2011. In addition, he is the CCNA and CCNP instructor and the coordinator of the Cisco Networking Academy Program since 2002 at this university. He worked between September 2017 and June 2018 as a visiting professor at the Department of Computing Science at the University of Alberta, Edmonton, Canada supported by TÜBİTAK-BİDEB 2219 Post-Doctoral Fellowship. He has many journal papers and international conference proceedings. he served as Associate Editor for the Journal of IEEE Access and the Turkish Journal of Electrical Engineering and Computer Science-TUBITAK from 2018 to 2021. He entered the 2% of the "World's Most Influential Scientists" list published by Stanford University researchers in 2020 and 2021. His current research interests include computer networks and security, cyber-security, software architectures, software testing, IoT/M2M applications, complex networks, graph visualization, knowledge discovery, and data fusion.