IJICTR
International Journal of Information & Communication Technology Research

Volume 14- Number 3 – 2022 (27 -36)

Ransomware Modeling Based on a Process Mining Approach

Ali Aghamohammadpour
Department of Computer Engineering,
Science and Research Branch, Islamic Azad University,
Tehran, Iran

Ebrahim Mahdipour*
Department of Computer Engineering,
Science and Research Branch, Islamic Azad University,
Tehran, Iran
mahdipour@srbiau.ac.ir

Iman Attarzadeh
Department of Computer Engineering, Faculty of Engineering,
Central Tehran Branch, Islamic Azad University,
Tehran, Iran

Received: 16 April 2022 – Revised: 6 June 2022 - Accepted: 17 August 2022

Abstract—**Ransomware attacks are taking advantage of the ongoing coronavirus pandemics and attacking the vulnerable systems in the health sector. Modeling ransomware attacks help to identify and simulate attacks against security environments, using likely adversary techniques. Process Mining (PM) is a field of study that focuses on analyzing process logs linked with the execution of the processes of a system to acquire insight into the variety of characteristics of how the functions behave. This paper presents a PM conformance-based approach to determining ransomware processes. First, frequent ransomware techniques were identified using state-of-the-art MITRE ATT&CK. Then, a model was developed to gather ransomware techniques using a process-based approach. The PM-based Prom tool is used to check the conformance of malware processes alongside the presented model to illustrate its efficiency. The model can identify chain processes associated with ransom-related behaviors. In this study, the presented model was evaluated using thirty common malwares in the healthcare industry. The approach demonstrates that this model could successfully classify ninety percent of malware instances as ransomware and non-ransomware. Finally, guidelines for future research are provided. We believe the proposed method will uncover behavioral models that will enable us to hunt ransomware threats.**

Keywords: *Process Mining; Ransomware Hunting; Threat Modeling; Threat Intelligence; Threat Hunting.*

Article type: Research Article

Publisher: ICT Research Institute

I. INTRODUCTION

In the last few years, ransomware has become a significant concern for many institutions, especially those in the healthcare industry. Many anti-virus products are ineffective against zero-day ransomware, resulting in a large amount of data loss. An effective way of profiling malicious software is through dynamic

* Corresponding Author

[DOI: 10.52547/itrc.14.3.27]

malware analysis [1]. Malware analysis tries to learn how the malware operates to develop efficient defenses.

Threat hunting is the process of determining system infection with malware and its exact behavior [2]. It refers to searching a network or endpoint for threats that are about to launch an attack or accomplish their objective despite security measures and has eluded security measures proactive. Iterative search processes detect and isolate advanced threats that circumvent existing security measures [3]. The most common data source for this type of research is event logs. PM makes data analysis possible in a timely and efficient manner.

The process mining research field is closely related to log and data mining. A process mining technique consists of three main components: process discovery, conformance checking, and enhancement [4]. Process discovery is the most common strategy in process mining, and it entails building a model by studying event logs obtained from systems. The second type of process mining analysis, conformance checking, compares existing process models against real event logs to see if the reported behavior matches the predicted [5]. In addition to improving current process models, enhancements further exploit process insights, such as performance analysis, by reconstructing new processes based on previous models.

Providing online process discovery and conformance checking could be one direction for researching current events [6]. As a result, process mining is an ideal approach for real-time analysis. Process mining techniques could affect hunting cyberattacks to prevent them before an imminent attack. Conformance checking activities evaluate by fitness. Measuring how well the model aligns with reality can be made by analyzing event logs and comparing them to process models, assuming that the event log contains acceptable or correct behavior [7], [8]. By considering that the process model includes correct or proper behavior, we may assess the fitness of a process model and an event log. The behavior of the event log that is unacceptable can also identify [8]. Identifying ransomware patterns and evaluating the new process models for threat hunting are the objective of this paper.

The ATT&CK Framework was created by MITRE Corporation to identify and categorize abusive behavior based on real observations [9]. It makes a knowledge base about the attacker's tactics, techniques, and procedures (TTP) as an organized information collection. This framework has evolved and can now serve as a comprehensive source of information about attacker methods, models, and mitigation strategies with its creation. A tactic refers to the steps taken by an adversary to achieve a deliberate goal. A technique describes how an adversary achieves a tactical objective.

This study looks into the possibility of applying ATT&CK to help with the systematic development and enhancement of behavioral models. This study uses process mining techniques to demonstrate a process-based hunting method using process behavior analysis to determine if it is or is not ransomware. We use conformance-checking to create and evaluate a process-mining-based method for recognizing ransom attacks. We analyze several malware types. The results show that we can identify and hunt ransomware between many malwares.

Section II contains the background information as well as the problem statement in this study. Section III shows and discusses preliminary findings on ransomware modeling. Section IV includes the conclusions and recommendations for future work.

II. Related Work

Research studies on process mining solutions for threat hunting discuss in this literature review. The PM approaches are rarely used in cybersecurity and haven't been applied widely to threat modeling. However, this area of research reveals that PM is a viable method for security use cases [10]. PM for cybersecurity has explored the strategy of exposing outliers in the process.

Researchers in [11], [12] introduced using process mining for anomaly detection in the event logs of information systems. Malware behavior modeling is a solution that allows knowledge to be extracted and represented in cybersecurity [13]. This information is helpful for predicting malevolent behavior based on previous discoveries. Modeling malicious behavior with realistic experiments is a novel tool [14]. Researchers in [15] proved process mining can help the current challenges in cybersecurity. The authors of [14] present a PM-based approach to studying smartphone malware detection. They find patterns in traces created by system calls performed by mobile applications. Modeling performs with PM approaches, with the premise that malicious conduct carries out by a series of system calls. Researchers in [15] used process mining to investigate attacks on a small application. The process model could observe in the annotated texts, which described the attack strategy of the participants. According to research in [16], insider attacks do not need to be intentional to succeed. Process discovery is to mimic the workplace environment for social engineering efforts.

According to [17], 42% of health delivery organizations had faced multiple ransom-related attacks in the previous two years, and 36% had faced a third-party assault. In this area, the other technique is to identify and inspect ransomware. In a comprehensive review of ransomware attackers, authors [18] examined recent studies, their essential contributions, and their limitations. Bharani [19] gathered event logs from harmless programs and ransomware families. Then, they could identify ransomware based on a process model they developed for each software. The essential factor of the process model is counting the number of iterations for each event. The event logs linked with a system's processes, which coordinate steps to achieve a goal, may be analyzed using PM techniques [20]. The process discovery techniques assist in identifying the most

[DOI: 10.52547/itrc.14.3.27]

appropriate models for describing the behavior inferred from the event logs. The Inductive Miner is a discovery algorithm dealing with extensive event records [21]. Several applications, such as the ProM toolkit [22], can automate the process of PM-based systems behavior analysis. To model ransom-related processes, we report the preliminary results of a research effort that aims to use PM techniques in this work.

Process discovery and conformance checking are two of the most prominent patterns in process mining. In some papers, these two motifs are integrated [23], [24]. Typically, a reference model generates the use of process discovery, and conformance checks apply to new instances; either process discovery [25] or conformance checking [26] can use.

MITRE ATT&CK is a knowledge base built on malicious tactics and techniques. ATT&CK entries for ransomware include publicly reported techniques and methods [27], MITRE ATT&CK ransomware list [27]. D3FEND is a MITRE complement to ATT&CK that includes numerous defense tactics [28]. Process analysis is a detection tactic in this framework. It includes monitoring a running application process and assessing it for particular behaviors or situations that indicate adversary activities. The accumulated knowledge of ATT&CK can use as a model of the attacker's behavior using PM techniques.

TABLE I. MITRE ATT&CK RANSOMWARE LIST [27]

Ransomware Name	MITRE ATT&CK ID
Bitpaymer	S0570
Conti	S0575
Egregor	S0554
JCry	S0389
LockerGoga	S0372
Maze	S0449
MegaCortex	S0576
Netwalker	S0457
NotPetya	S0368
Pay2key	S0556
Pysa	S0583
Ragnar Locker	S0481
Revil	S0496
RobbinHood	S0400
Ryuk	S0446
SamSam	S0370
SynAck	S0242
Thief quest	S0595
WannaCry	S0366
Xbash	S0341

It has been suggested in [29] that many pipeline elements for data processing can be automated using the proposed architecture, including learning alert templates, splitting an alert graph into individual incidents, and using factor graphs to rank these incidents. MITRE ATT&CK also incorporates several phases in its design. CSA published Cybersecurity Risk Assessment for Critical Information Infrastructure [30]. MITRE ATT&CK framework is an example of threat modeling to trace the threat sequence.

Kestrel language model was introduced by authors in [31]. In the cyber threat hunting process, Kestrel adds an element of abstraction, reducing repetition. According to the authors of [32], enterpriseLang is a MAL-based domain-specific language built using design science principles. A business system can evaluate its overall cyber security Using the language. EnterpriseLang, supplemented by additional data source, covers most enterprise-level threats.

Process mining has been studied very little in cybersecurity. Many potential applications have yet to investigate. Previous works demonstrate process mining can hunt threats and evaluate simulated data successfully. According to this research, process mining as a tool for threat hunting is an exciting area, and further investigation of process mining may be in order. It highlights the importance of accurate modeling in standard hunting procedures.

III. RESEARCH APPROACH

Samples from the same malware family would behave very similarly [33]. Structured behavior, such as a process-like activity consisting of multiple sequential events, can be used to identify adversary threats. The study aims to identify new ransomware variants based on their behavior. Comparing the conformance of new malware with our unique process-based model will uncover new ransomware variants.

MITRE ATT&CK is a total knowledge base for malware techniques and tactics. Tactics represent the adversary's tactical goal and the reason for acting. The MITRE ATT&CK Matrix for Enterprise consists of 14 tactics. Techniques represent how an adversary gain a tactical goal by acting. For example, an adversary may dump credentials to achieve credential access (T1003), as described in the MITRE ATT&CK knowledge base [34].

MITRE had specified the list of widespread ransomwares, MITRE ATT&CK ransomware list [27]. We focused our technical studies based on this list to present our model. The MITRE created a ransomware activity heat map of frequent ransomware tactics and techniques [27]. In this heat map, each technique has a score from 1 to 20 based on its frequency of occurrence in ransomware groups. Techniques with a score of less than three were excluded to increase the accuracy in making the heat map of ransomware activity.

Ransomware frequent activity heat map [27] presents ransomware threat group's frequent techniques based on open-source reports that are not limited to ATT&CK pieces [27]. The heading row in Ransomware frequent activity heat map [27] represents the tactics, and each id describes the techniques. Depending on an attacker's high-level goal and the malware environment, each ransomware has a unique set of circumstances.

[DOI: 10.52547/itrc.14.3.27]

TABLE II. RANSOMWARE FREQUENT ACTIVITY HEAT MAP [27]

Initial Access	Execution	Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery	Lateral Movement	Collection	Command and Control	Exfiltration	Impact
T1190	T1059	T1547	T1547	T1564	T1110	T1135	T1021	T1005	T1105	T1041	T1486
T1566	T1129	T1543	T1543	T1562	T1003	T1057					T1490
T1078	T1569	T1078	T1055	T1070		T1082					T1489
			T1078	T1036		T1016					
				T1112							
				T1027							
				T1055							
				T1218							
				T1078							

In this study, ransom-related processes were extracted by joining techniques based on various technical reports of analysis.

The Rapid7 Endpoint detection rules cover various MITRE ATT&CK techniques [35]. There are possible frequent processes in behavioral analysis reports of each ransomware. To reduce complexity, we considered each process contains two simple steps with a source and a destination technique.

To develop a ransomware attack model 122 ransom-related processes were collected based on discussed frequent techniques and reviewing technical analysis of ransomwares in MITRE ATT&CK ransomware list [27] [35]- [53]. Each process contains a specified source and destination technique id (TID), The collected ransomware processes. For example, a ransomware may attempt to disable backup software in case 44. T1489 is the source, and T1562 is the destination technique. Inductive miner is a method for constructing a log's process tree [21]. In this study, ProM was employed to identify a model from collected ransomware processes using an inductive miner approach.

A unique model was developed to infer ransomware attack technique association. A directly follows visual miner (DFvM), an extension of the inductive visual miner (IvM) used as an easy-to-use process mining exploration tool, Process tree of frequent techniques in ransomware family using Inductive miner

.

DFvM performs process discovery automatically and iteratively to discover the process models and visualize the different paths of process cases. Here, the green and red circles indicate the start and end of the process, respectively. Process tree of frequent techniques in ransomware family using Inductive miner

shows the DFvM model discovery of all 122 process cases. In the beginning, the path split into several directions. The number seven in T1078 tells us that it engaged in 7 cases and five processes begin with T1078, while T1190 engaged in four cases. Furthermore, it can see that there is one possible path between T1190 and T1078 (CASE 89).

TABLE III. THE COLLECTED RANSOMWARE PROCESSES

Case	Source TID	Destination TID	Reference
CASE 1	T1003	T1003	[35]
CASE 2	T1003	T1055	[35]
CASE 3	T1003	T1059	[35]
CASE 4	T1003	T1110	[35]
CASE 5	T1003	T1218	[35]
CASE 6	T1021	T1021	[35]
CASE 7	T1021	T1059	[35]
CASE 8	T1021	T1562	[35]
CASE 9	T1021	T1569	[35]
CASE 10	T1027	T1027	[35]
CASE 11	T1027	T1059	[35]
CASE 12	T1027	T1566	[35]
CASE 13	T1036	T1036	[35]
CASE 14	T1036	T1055	[35]
CASE 15	T1036	T1059	[35]
CASE 16	T1036	T1218	[35]
CASE 17	T1036	T1564	[35]
CASE 18	T1036	T1566	[35]
CASE 19	T1055	T1055	[35]
CASE 20	T1055	T1059	[35]
CASE 21	T1055	T1218	[35]
CASE 22	T1055	T1569	[35]
CASE 23	T1059	T1059	[35]
CASE 24	T1059	T1070	[35]
CASE 25	T1059	T1082	[35]
CASE 26	T1059	T1105	[35]
CASE 27	T1059	T1110	[35]
CASE 28	T1059	T1112	[35]

[DOI: 10.52547/itrc.14.3.27]

CASE 29	T1059	T1218	[35]
CASE 30	T1059	T1547	[35]
CASE 31	T1059	T1562	[35]
CASE 32	T1059	T1564	[35]
CASE 33	T1059	T1566	[35]
CASE 34	T1070	T1070	[35]
CASE 35	T1070	T1486	[35]
CASE 36	T1105	T1218	[35]
CASE 37	T1105	T1566	[35]
CASE 38	T1110	T1110	[35]
CASE 39	T1112	T1547	[35]
CASE 40	T1112	T1562	[35]
CASE 41	T1112	T1564	[35]
CASE 42	T1218	T1218	[35]
CASE 43	T1218	T1566	[35]
CASE 44	T1489	T1562	[35]
CASE 45	T1490	T1562	[35]
CASE 46	T1543	T1543	[35]
CASE 47	T1543	T1564	[35]
CASE 48	T1562	T1562	[35]
CASE 49	T1564	T1564	[35]
CASE 50	T1566	T1566	[35]
CASE 51	T1569	T1569	[35]
CASE 52	T1027	T1057	[36]
CASE 53	T1036	T1057	[36]
CASE 54	T1055	T1489	[36]
CASE 55	T1055	T1082	[36]
CASE 56	T1057	T1055	[36]
CASE 57	T1059	T1486	[36]
CASE 58	T1070	T1027	[36]
CASE 59	T1070	T1490	[36]
CASE 60	T1486	T1562	[36]
CASE 61	T1489	T1016	[36]
CASE 62	T1490	T1489	[36]
CASE 63	T1490	T1027	[36]
CASE 64	T1490	T1036	[36]
CASE 65	T1547	T1059	[36]
CASE 66	T1562	T1070	[36]
CASE 67	T1562	T1490	[36]
CASE 68	T1021	T1218	[37]
CASE 69	T1055	T1547	[37]
CASE 70	T1218	T1489	[37]
CASE 71	T1489	T1490	[37]
CASE 72	T1547	T1021	[37]
CASE 73	T1562	T1055	[37]
CASE 74	T1003	T1057	[38]
CASE 75	T1021	T1003	[38]
CASE 76	T1057	T1041	[38]
CASE 77	T1059	T1021	[38]
CASE 78	T1078	T1218	[38]
CASE 79	T1218	T1059	[38]
CASE 80	T1486	T1218	[38]
CASE 81	T1003	T1105	[39]
CASE 82	T1027	T1036	[39]
CASE 83	T1041	T1486	[39]
CASE 84	T1105	T1059	[39]
CASE 85	T1112	T1486	[39]
CASE 86	T1486	T1027	[39]
CASE 87	T1562	T1547	[39]
CASE 88	T1055	T1486	[40]
CASE 89	T1190	T1078	[40]
CASE 90	T1190	T1016	[40]
CASE 91	T1055	T1027	[41]
CASE 92	T1078	T1110	[41]
CASE 93	T1135	T1021	[41]
CASE 94	T1566	T1110	[41]
CASE 95	T1078	T1569	[42]
CASE 96	T1078	T1059	[42]
CASE 97	T1190	T1059	[42]
CASE 98	T1190	T1569	[42]
CASE 99	T1566	T1569	[42]
CASE 100	T1055	T1021	[43]
CASE 101	T1486	T1569	[43]
CASE 102	T1566	T1059	[43]
CASE 103	T1569	T1547	[43]
CASE 104	T1569	T1543	[44]
CASE 105	T1021	T1486	[45]
CASE 106	T1059	T1569	[45]
CASE 107	T1082	T1021	[45]
CASE 108	T1569	T1059	[45]
CASE 109	T1059	T1129	[46]
CASE 110	T1055	T1078	[47]
CASE 111	T1078	T1021	[47]
CASE 112	T1005	T1486	[48]
CASE 113	T1036	T1070	[49]
CASE 114	T1486	T1112	[49]
CASE 115	T1562	T1036	[49]
CASE 116	T1490	T1486	[50]
CASE 117	T1543	T1490	[50]
CASE 118	T1055	T1135	[51]
CASE 119	T1135	T1490	[51]
CASE 120	T1490	T1082	[51]
CASE 121	T1105	T1486	[52]
CASE 122	T1005	T1041	[53]

[DOI: 10.52547/itrc.14.3.27]

Volume 14- Number 3 – 2022 (27 -36)

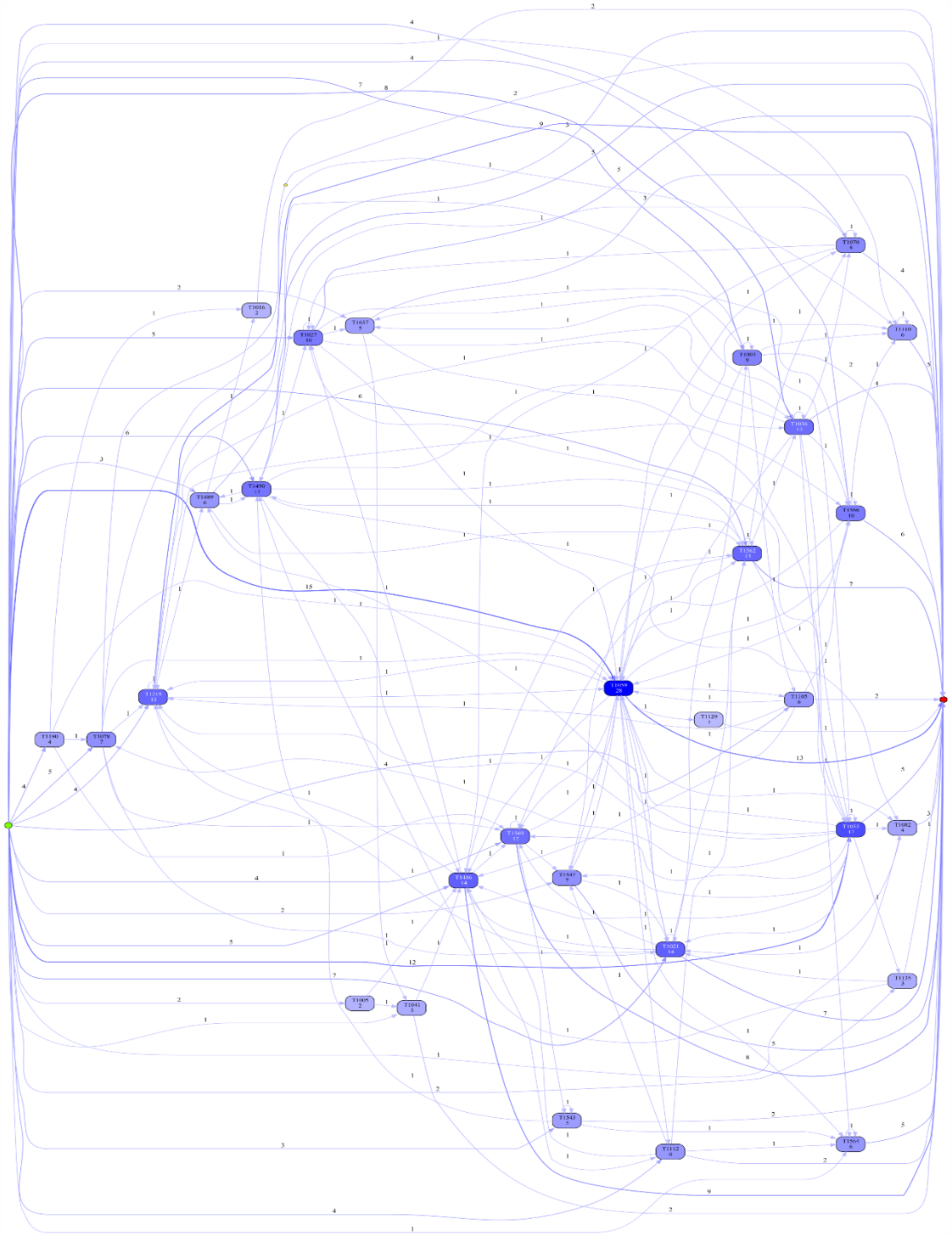


Fig. 1. Process tree of frequent techniques in ransomware family using Inductive miner

IV. RESULTS AND DISCUSSION

A detailed timeline of malware attacks is difficult to determine. To understand the events, we can compare the conformance of malware techniques with our model. The proposed method describes the matches between the log traces and the model based on the deviation of event log traces from appropriate process models.

To generate a malware techniques list, we analyzed thirty common malware in the healthcare industry [54] using AlienVault Open Threat Intelligence [55]. We collected frequent ransomware processes in each malware to mine each process cases. As a sample, The Conti malware frequent ransom-related process cases shows analyzed frequent ransomware techniques of the Conti malware. Using the ProM, we compare the results of the conformance check.

A process tree can be directly transformed into a Petri Net. A Petri Net is a graph model for the control behavior of systems

[DOI: 10.52547/itrc.14.3.27]

exhibiting concurrency in their operation [5]. Concerning Petri net token firing rule if there is no token for a trace to replay, the artificial token is placed for trace parsing. In the end, artificial and left tokens were considered for mismatch measurement.

Volume 14- Number 3 – 2022 (27 -36)

TABLE IV. THE CONTI MALWARE FREQUENT RANSOM-RELATED PROCESS CASES

Case	Source TID	Destination TID
CASE 1	T1021	T1021
CASE 2	T1021	T1059
CASE 3	T1027	T1027
CASE 4	T1027	T1059
CASE 5	T1027	T1566
CASE 6	T1055	T1055
CASE 7	T1055	T1059
CASE 8	T1059	T1059
CASE 9	T1059	T1110
CASE 10	T1059	T1566
CASE 11	T1110	T1110
CASE 12	T1566	T1566
CASE 13	T1027	T1057
CASE 14	T1055	T1489
CASE 15	T1057	T1055
CASE 16	T1059	T1486
CASE 17	T1489	T1016
CASE 18	T1490	T1489
CASE 19	T1490	T1027
CASE 20	T1489	T1490
CASE 21	T1059	T1021
CASE 22	T1486	T1027
CASE 23	T1055	T1486
CASE 24	T1055	T1027
CASE 25	T1078	T1110
CASE 26	T1135	T1021
CASE 27	T1566	T1110
CASE 28	T1078	T1059
CASE 29	T1055	T1021
CASE 30	T1566	T1059
CASE 31	T1021	T1486
CASE 32	T1055	T1078
CASE 33	T1078	T1021
CASE 34	T1490	T1486
CASE 35	T1055	T1135
CASE 36	T1135	T1490

The ProM has a plugin that can mine a Petri net directly using the Inductive Miner technique. In this study, the ProM tool was used to discover a Petri net model from an event log with the Inductive Miner technique.

The Replay a Log on Petri Net plugin in ProM used to analyze the pre-processed cases and petri-net model of expected behavior, Conti trojan *replay result projected with alignments*

Replay results visualize using project alignment to log once they obtain. The focus is on the following statistics: Trace fitness, move-log fitness, and move-model fitness, and we use them for the classification task. The move-log fitness computes the trace of the event log on the process model used in the conformance checking activity and lets you see where it differs from the process model. The move-model fitness shows where the model differs from the event log. Trace fitness offers the overall fitness of the model and log, which considers move-log and move-model fitness. Trace fitness has a broad number of 0 to 1 relative to the process. Formerly, once the model could replay the traces completely, the function returns "1," otherwise "0."

Fitness quantifies, by which the process discovery model can accurately express all behaviors recorded in the event log. The final fitness score is calculated in equation (1) [56].

$$f(\sigma, N) = \frac{1}{2}\left(1 - \frac{\sum_{i=1}^{k} n_i m_i}{\sum_{i=1}^{k} n_i c_i}\right) + \frac{1}{2}\left(1 - \frac{\sum_{i=1}^{k} n_i r_i}{\sum_{i=1}^{k} n_i p_i}\right) \quad (1)$$

Assume L is an event log and N is a WF-net (a Petri net with a start and an end place). Note that σ ∈L is an event sequence of L, ∑ denotes the sum of all produced, consumed, missing, and remaining tokens, and applies the same formula. Let pN,σ indicate the number of created tokens when replaying σ on N. mN,σ is the alternative duplicate task that is never repeated together in one sequence number of missing tokens when replaying σ on N. cN,σ is the number of consumed tokens. rN,σ is the number of remaining tokens. If fitness is 1, the discovery process model can replay all traces in the event log.

[DOI: 10.52547/itrc.14.3.27]

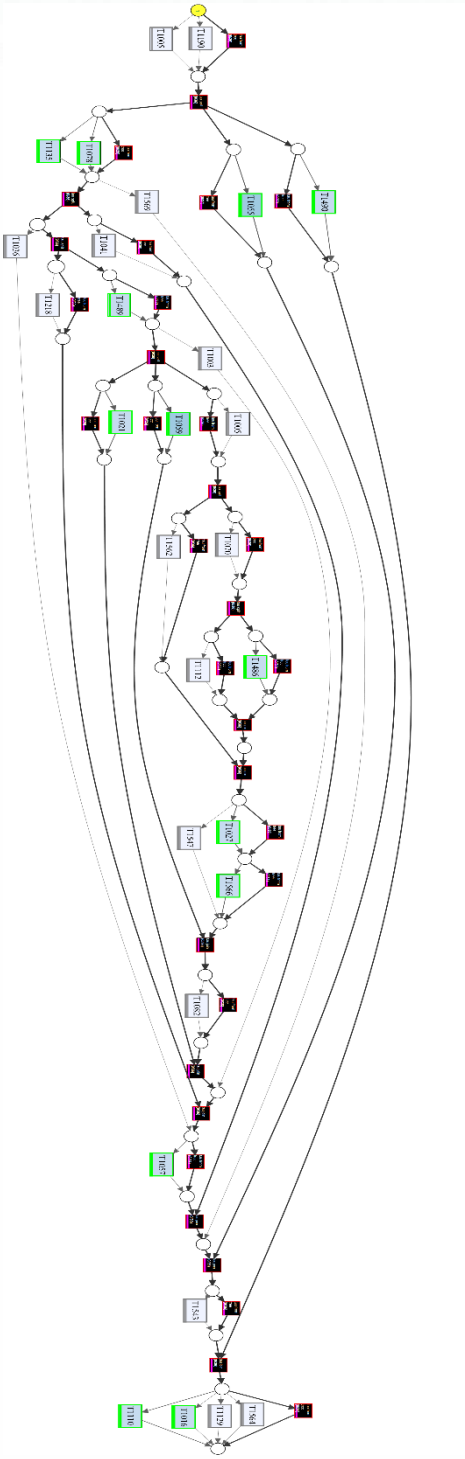


Fig. 2. *Conti trojan replay result projected with alignments*

This study's conclusions apply just to the ransomware family classification and not to the malware detection, which presents a fundamentally different challenge as a set of known malwares may use some ransom-related techniques.

This procedure was applied to all thirty malware samples to compare the conformance of each sample trace. Table 5 illustrates the results of our analysis to measure the trace fitness of thirty different malware samples. The left axis shows the trace fitness of malware processes regarding the presented ransomware frequent techniques model.

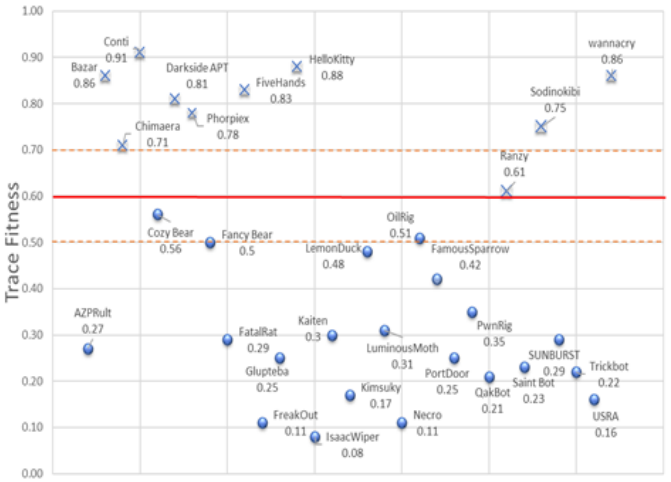
We find that when the trace fitness of samples is more than 0.6, we could classify malware as close-fitting to our model, and w hen the trace fitness of samples is less than 0.6, we could classify them as non-ransomware malware. As shown in Table 5, the trace fitness value of ransomware samples is more than 0.6.

According to the results of Table 5, ten malware's trace fitting is more than 0.6. Also, there are twenty malware's trace fitting is lower than 0.6. It is straightforward to adjust the threshold value to set the desired balance between false positives and negatives. A rigid threshold setting is not a simple task and allows suspicious patterns to pass through the presented classification mechanism. If the trace fitness is less than the threshold value, the file flag is non-ransomware; this implies the need to look for patterns with an optimal length for malware identification by setting the threshold.

Based on our experiments, a ten percent threshold can be considered a fair tradeoff. If the trace fitness value is more than 0.7, the instance classifies as ransomware, and if the trace fitness value is less than 0.5 as non-ransomware. With a threshold of 10 percent, the model can classify around 90% of ransomware instances but misclassify almost 10% of the malware instances as non-ransomware.

Our experiments demonstrate that conformance checking can identify and hunt ransom-related malware with their behaviors. In contrast to previous work [57], we found technique-based conformance checking to be a better method for hunting ransomware. Whenever we encounter several unknown files, we check their conformance with the presented model to determine the hunting hypothesis and investigate those files further.

TABLE V. FITNESS OF THE MALWARE INSTANCES WITH THE PRESENTED MODEL.



In threat hunting approach, the challenge is to reduce the time taken to classify the file; using the threshold, we filter out the “definitely ransomware behave” and the “definitely non-ransomware behave” malwares leaving a small number of files in the gray area that need to be evaluated with the dueling threat hunting approach.

The presented advanced method can hunt new ransomware variants. This method can alert the hunter to potential threats in a threat hunting system. The solution doesn’t require a signature database and provides classified data on ransomware and other types of malwares.

[DOI: 10.52547/itrc.14.3.27]

V. CONCLUSION

The coronavirus pandemic has caused a considerable growth in the use of online technologies to support remote work, resulting in a sharp rise in ransomware crimes across the globe.

Malware variants with different behaviors may escape detection by presenting unusual behaviors for newly collected samples. In addition, they may intend to misdirect detection and classification systems by mimicking a similar behavior found in another ransomware.

We explored process mining usage for ransomware hunting. Our first step was to collect analysis reports of selected ransomware groups and extract their process attributes. For ransomware hunting, we identified the set of processes to be employed. Additionally, we propose a novel ransomware process model. Next, we assessed the model's accuracy. The model enabled us to identify modified versions of ransomware samples. Our model could successfully classify ransomware and non-ransomware malware with ninety percent accuracy. Throughout our research, we have identified several potential areas for future research. Pre-processing in our ransomware hunting method relies on the most common techniques found in resources. The future work of this study will be to compare our approach with more hunting methods, using attack datasets containing campaign attacks. As we move forward, we plan to have more ransomware samples and conduct experiments with other process mining algorithms, including enhancements. Also, we can apply this method to classifying other malware families.**Error! Reference source not found.**

Ali Aghamohammadpour received his M.Sc. degree in Computer Science from Shahid Beheshti University, Tehran, Iran, in 2015. He is a Ph.D. student specializing in Computer Architecture; Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran. His research interests include Cybersecurity, System Architecting, and Data Engineering.

Ebrahim Mahdipour received a B.Sc. degree in Computer Engineering, specializing in hardware engineering, from the Dezful Branch, Islamic Azad University, Dezful, Iran, in 2003. He received his M.Sc. degree in Computer Engineering, specializing in Computer Architecture in 2006, and the Ph.D. degree in Computer Engineering, specializing in Computer Architecture, from the Science and Research Branch, Islamic Azad University, Tehran, Iran, in 2012. He is the Founder and Director of the Cybersecurity Research Center and CERT-IAU. He is currently an Assistance Professor with the Department of Computer Engineering, Science and Research Branch, Islamic Azad University. His research interests include cyber security, blockchain, and big data.

Iman Attarzadeh received his B.Sc. in Computer Engineering in 2002, his M.Sc. in Computer Architecture in 2005, and his Ph.D. in Software Engineering from the University of Malaya (UM), Kuala Lumpur, Malaysia, in 2011. He is an Assistant Professor in the Department of Computer Engineering, Faculty of Engineering, Central Tehran Branch, Islamic Azad University, Tehran, Iran. His research focuses on Data Science, Data Analytics, Pattern Recognition, Software, and Systems Engineering, Machine Vision.

[DOI: 10.52547/itrc.14.3.27]