A Big Data Fusion to Profile CPS Security Threats Against Operational Technology

Karl Biron[‡], Wael Bazzaza[‡], Khalid Yaqoob[*], Amjad Gawanmeh[‡], Claude Fachkha[‡]
[*]Dubai Electronic Security Center (DESC), UAE
[‡]University of Dubai, College of Engineering and IT, Dubai, UAE.
kvbiron@ud.ac.ae, mwael@ud.ac.ae, khalid.yaqoob@desc.gov.ae, agawanmeh@ud.ac.ae, cfachkha@ud.ac.ae

***Abstract*—Internet security measurements are fundamental techniques to detect cyber attacks and generate intelligence. However, such methods are limited in terms of relevancy, scalability, and data availability when it comes to Operational Technology (OT). Therefore, in this paper, we build cyber security capabilities to collect, detect, analyze, and visualize, in near real-time, cyber-attacks targeting Cyber-Physical Systems (CPS). The latter is a critical component for Industry 4.0 and smart technologies. In order to achieve our tasks, we propose a big data fusion model, which correlates among two trap-based monitoring systems, namely, darknet and honeypot. With approximately hundred deployed sensors (monitors) over a six-month period, we have been able to collect several unauthorized and malicious activities originating from 226 countries. Furthermore, our investigation has revealed various scanning strategies such as CPS-focused scans, in addition to real exploits used by external sources to infiltrate our network. Finally, this study highlighted the importance of such monitoring systems and compare the efficiency among them with an aim to complement existing CPS and on-premise OT security solutions.**

***Index Terms*—Honeypot, Darknet, CPS, ICS, OT, SCADA, Security**

I. Introduction

Cyber security threats and attacks have been on the rise during the past decade. The advancement of technology in cyber-physical systems (CPS) has led to many advantages but also created several new advanced threats. A real example would be the event known as the Estonia Blackout, which involved a series of cyber-attacks directed towards government entities such as ministries, financial institutions, Internet Service Providers (ISPs), and telecommunication land line systems [1]. These attacks that targeted critical infrastructures caused a so-called "cyber blackout", which effectively took down various online services including critical ones such as banking. Attack methods range from denial of service with ping floods to the usage of botnets to distribute spam at a high magnitude and intensity. As a result, citizens were cut-off from news due to broadcasters not being able to deliver news and specifically, government employees were not able to communicate by either email or telephones. In 2017, Triton malware was used to attack the petrochemical plant in the Kingdom of Saudi Arabia (KSA) and caused it to shutdown to prevent an explosion [2].

While CPS is included in several critical sectors, such as power, nuclear plants, and urban infrastructures, these systems are often subtle to various cyber security threats due to several reasons. First, these systems were not designed in the first place to be part of a public network, i,e, the Internet, they were in fact built and operated in a segregated and contained environment, this has changed over the past two centuries due to the need for remote operations as well as providing live and continuous support from various vendors. In addition, these CPS are often equipped with IoT devices, many of these are low cost and low-performance devices that are designed with built-in security features. As a result, several security concerns emerged about these critical systems.

Therefore, cyber-attacks of this large-scale and intensity have called for further research and development efforts in beefing up cyber security measures to detect, analyze, and even prevent potential attacks. This paper elaborates on building capabilities to achieve such mitigation aims. In this context, we propose a data-driven fusion approach that uses trap-based monitoring systems, namely, darknet and honeypot. The proposed detection approach is expected to yield important findings regarding common behavior, trends, similarities among CPS threat sources.

The Darknet is commonly misinterpreted as the Dark Web which is an entirely different affair in the broad field of information technology and information systems. A better and more accurate term for Darknet would be a network telescope [3]. A Darknet is a segment of an assigned IP Address block that has no active services running. The traffic that arrives in this IP space is typically unwanted and thus should arise suspicion of malicious intent or activity. A Honeypot however is an information security terminology that pertains to a mechanism that is intended to counteract, deflect and detect directed efforts of unauthorized usage of systems. In a general sense, a honeypot contains data or a set of data that may seem to appear as a legitimate component of a network or system but in reality is actually a standalone system that is monitored. It lures potential attackers as these honeypots may seem to contain resources or information that could be of value to attackers. A comparative study and more about a trap-based monitoring system can be found in [3].

As part of our contribution, we provide a big data framework for collecting data about various types of attacks from two types of monitoring systems: honeypot and darknet. Then, the data is analyzed for both sources, and a comparative analysis between them is performed. This can provide insights and recommendations on which one is more efficient in terms of sensing attacks and identifying potential threats. In addition, this will add capabilities to identify potential CPS threats and provide defense measures against most common attacks. As a result, feedback can then be provided to potential CPS entities on enhancing their security measure against such potential and more common attacks.

II. Related Work

This section presents a summary of several contributions in the fields of CPS. The work in [4] used monitoring systems to identify suspicious network flows. Moreover, the authors in [5] made use of the Industrial Control System (ICS) HoneyPhy capable of constructing honeypots for complex CPS. The ICS honeypot used in their project was based on open-source technology, namely, Conpot [6]. The authors in [7] and [8] provided analysis on most common attacks against different types of protocols used in CPS in general, and Modbus in particular. The authors in [9] proposed a cyber-physical honeypot intended for industrial control networks. In addition, the authors in [10] proposed another honeypot using virtual and real programmable logic controllers that interact with a physical process model. Another physical-aware CPS honeypot framework was proposed by Litchfield [11] which was deployed on industrial control system with PLC to emulate simple DoS attacks.

In this work, we extend the contribution in [4] to add darknet analysis and correlate among various types of trap-based monitoring systems, instead of only one. Furthermore, we do not only correlate open-source tool as in [5], but also investigate darknet data. To the best of our knowledge, this is the first large (Internet-wide) study which investigates and correlates CPS analytics from both darknet and honeypot Big data perspectives.

III. Proposed Big Data Fusion Approach

The proposed approach involves running two trap-based monitoring systems, namely, honeypot and darknet. Such detection systems are used to collect different types of security analytics. However, we have tailored these systems to collect CPS-related data. First, the darknet, also known as unsolicited network traffic or network telescope, is composed of unused IP addresses. Therefore, any network traffic targeting such IP address space is considered unauthorized/unsolicited and suspicious. Darknet data is composed of one-way traffic, which does not communicate with third parties. In other words, such monitoring systems are run in passive mode. In addition to identifying misconfiguration, darknet data can be used to infer scanning activities, DDoS attacks, among others [12]. Second, in contrast to darknet systems, honeypots are interactive traps. As such, such systems can collect two-way communications. Compared to darknet, honeypot is considered richer in terms of intelligence gathering. However, it is also considered less secure. The reason behind this is that the more you interact with a third party, the more exposed you are [3].

In this paper, we develop our own (in-house) private cloud darknet system, which is a combination of 45 unused IP address space provided from our ISP provider, a router and a firewall extracting dedicated network traffic and feeds extractor. As far as the honeypot system is concerned, we used Conpot, an open-source SCADA honeypot [6] on another 45 IP addresses. Conpot is considered a low-interactive honeypot simulating programmable logic controllers (PLCs) and tracking several services such as Modbus, IPMI, S7Comm, BACnet, among others. Such services are used in critical infrastructure networks such as building automation, elevators, power and nuclear plants, etc.

On one hand, the darknet system is designed to collect data using deep packet inspection and detection tools in a raw PCAP format. The collected network data is then filtered, sanitized, and converted tabulated files for big data processing. The script filters internal traffic noise such as ARP from the PCAP and then converts it into parsed files with geographical data that is extracted from open-source tools (e.g., whois). On another hand, the honeypot component tracks SCADA communications such as traffic targeting ICS. Potential threats that attempt to establish communication with the honeypot sensors are being monitored and analyzed in an automated fashion. The honeypot system produces results in the form of a log file that has verbose information regarding all communication attempts toward the sensor. Similarly, such information are then filtered, sanitized, and converted into tabulated files for big data processing.

All the scripts use open-source tools to append additional geographical data for further analysis. Among the analytics, we extract the SCADA targeted ports, dates, time, source IP addresses, countries, cities, map coordinates, ISP that owns the source IP addresses, ASN numbers, among others.

Figure 1 depicts an overview of our system, which consists of several tools including Apache Zeppelin, Hadoop, Hive, Python, Spark, PHP, HTML, CSS, and JavaScript, which are used to process, analyze, and visualize the data generated from around 90 trap-based monitoring systems. First, datasets are aggregated from both Darknet and Honeypot networks for processing. Apache Zeppelin provides the back-end environment for the big data programming languages to interact while the data files are also concurrently being stored as Hadoop DFS for backup purposes. Consequently, Spark is used to compute the values using aggregation and other big data techniques while Python scripts are utilized to parse data into a format acceptable by HighCharts JavaScript. During this process, HiveQL queries generate tables, CSV files and Zeppelin based graphs. The results are gathered in processed files in specific directories. The files are retrieved by PHP codes and then submitted to the web development folder. HTML, CSS and JavaScript files are used together to create the visualization dashboard. Finally, accumulated analytics are then correlated and sent to a host which handles big data analytics and visualizations.

A. Data Acquisition and Processing

Our trap-based systems (darknet and honeypot) gather network malicious activities. On one hand, the Honeypot utilizes Conpot, an open-source tool specifically configured to emulate CPS and SCADA ports. The aim of this honeypot is to trick adversaries who are looking for real critical infrastructure vulnerabilities. Conpot has the capability to uncover malicious CPS network behaviors from session initiation to message transfer. On another hand, the darknet monitors are set into passive mode; an IP sinkhole. Darknet does not communicate

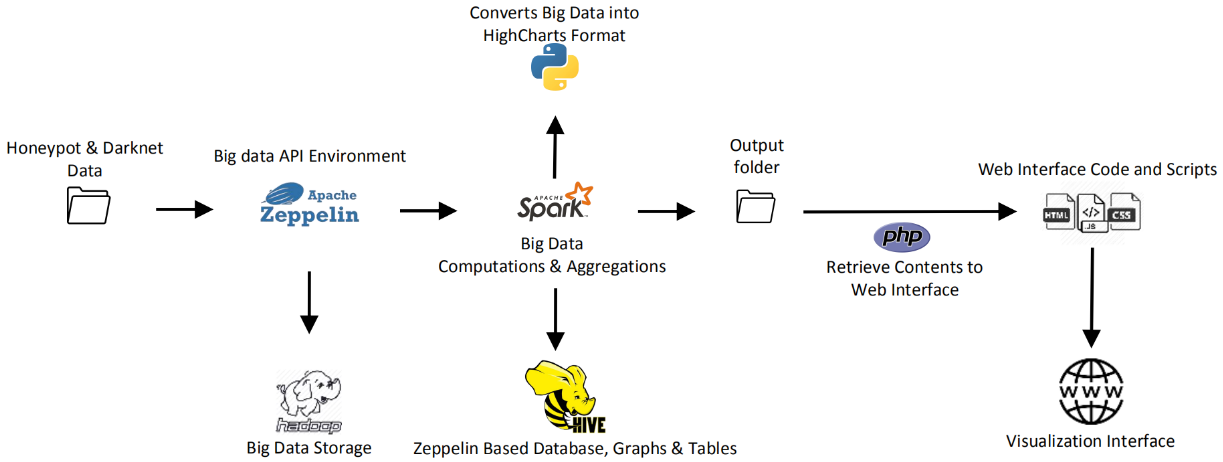
Converts Big Data into HighCharts Format
Honeypot & Darknet Data
Big data API Environment
Apache Zeppelin
Spark
Big Data Computations & Aggregations
Output folder
Web Interface Code and Scripts
php
Retrieve Contents to Web Interface
hadoop
Big Data Storage
HIVE
Zeppelin Based Database, Graphs & Tables
Visualization Interface

Fig. 1. Big Data Storage, Processing and Visualization System

with adversaries. However, it collects the first initiation packets in communications from all types of scanning (TCP SYN, X-MAS, etc.) to reply back-scattered (e.g., ACK) from DDoS victims, among others [3].

The Honeypot results are acquired and processed in near real-time. A script is used to parse the monitoring logs. First, the verbose logs are broken down into separate fields specifically extracted and organized by the script. Second, the script creates records related to traffic (network flow) with information such as SCADA port, date, time, source IP address, country, region, latitude, longitude, ISP, ASN, severity, and sensor-ID. Finally, the script then pushes the collected records into a single file (in CSV or JSON format). This file is then appended to a second bulk file as an update. This bulk file contains all historical data and is constantly being updated by all the Honeypot sensors. In addition, this bulk file is being utilized by the big data and visualization side to process and convey statistical and graphical representation of the data.

Darknet sensors produce PCAP files that contain network traffic logs. A couple of scripts are used to process such files. First, the PCAP file is filtered of any misconfiguration traffic which consists primarily of internal traffic and device broadcasts. Specifically, traffic with ARP, RARP and destination hosts FF02::FB, FF02::1 and IP address 0.0.0.0. In a nutshell, any non-routable meta-address used to designate an unknown, invalid, or non-applicable target is filtered out. As such, the remaining probes and network activities are kept for processing and examination. Subsequently, we extract specific packet columns such as the date, time, source IP address, frame length, protocol, source port, destination port, country, region, latitude, longitude, ISP, ASN number, and sensor-ID to add more details to our analytics. Finally, the script then correlates the extracted information into a single file. Similarly to the Honeypot process, this file is then appended to a bulk file as an update.

The aim of the data acquisition and processing is to provide the big data and visualization segment of the project sanitized and organized results for further processing. The result from the process can provide insights that could lead to conclusive evidence of malicious behavior.

IV. Preliminary Results

In this section, the accumulated datasets from both monitoring systems are processed and analyzed to identify potential attack patterns and behavior. We start by profiling the most generic information, namely, the targeted ports and services.

A. CPS Data Overview: Darknet and Honeypot

Darknet is relatively easier to deploy when compared to honeypot. The latter requires a communication protocol configuration to interact with adversaries, whereas darknet is set in passive mode (solely listening). Consequently, darknet is considered more secure than honeypot, which can be detected by hackers through communication.

Darknet monitors have gathered data from 41,830 unique IP addresses while the Honeypot monitors have gathered traffic from 2,686 unique IP addresses. In our setup, Darknet is designed to collect all traffic (conventional and CPS. Whereas our honeypot monitors are designed to capture CPS data only. As such, the number of data collected from the darknet systems is bigger. The distribution of the collected data over half of 2019 is shown in Figure 2. On one hand, on a monthly average, 583 non-unique IP addresses have been found generating 97,430 packets over our CPS honeypot monitors. This is shown in the blue and grey boxes of Figure 2, respectively. On another hand, an average of 207,045 non-unique IP addresses has been found sending 4,451,990 packets over our darknet monitors on a monthly average. This is shown in the orange and yellow boxes of Figure 2, respectively. This analysis reveals that honeypot targeted threats are found to be more stable than darknet.

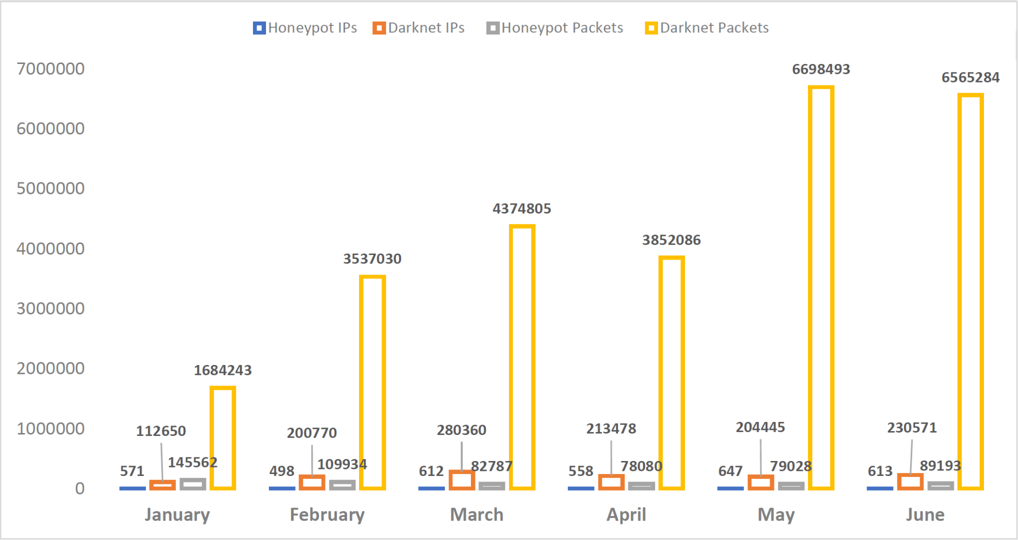

Fig. 2. Distribution over 6 Month Period in 2019 - Number of IPs and Packets

Furthermore, in order to provide an overview related to the source of threat. We geolocate the source IP addresses communicating to our monitors. Figure 3 illustrates the activities generated from the top continents targeting our honeypot and darknet sensors. For darknet data, almost half with 47.55% of information originated from Asia, 33.91% from Europe, 16.16% from North America, and the remaining minority from South America, Oceania, and Africa. For honeypot data, the results are not similar, more than half with 51.76% of the data originated from Europe, 32.19% from North America, 15.36% from Asia, and the remaining minority from South America, Oceania, and Africa. It is noteworthy to mention that in total, for both monitoring systems, the majority of the data is originated from Europe and the minority from Oceania.

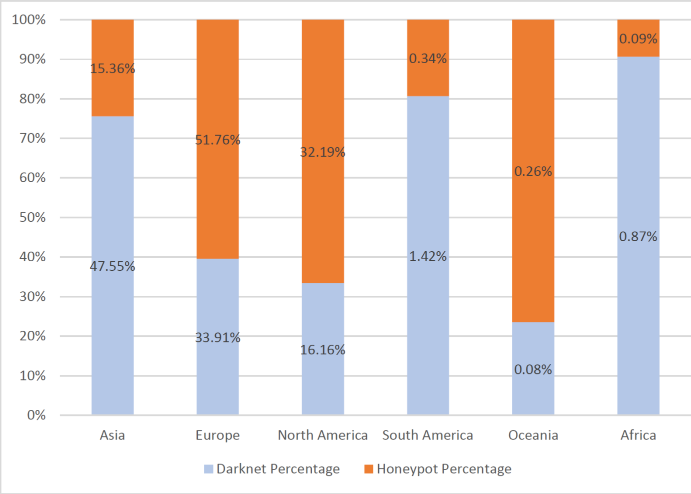

Fig. 3. Distribution of Data per Continent - Honeypot vs Darknet

The overlap between the darknet and honeypot datasets is 674 IP addresses as shown in Figure 4. This image allows us to identify the following scenarios: 1) dedicated, also known as focused, attacks targeting either honeypots or darknet only; 2) random attacks targeting any of the monitors; and 3) Internet-wide attacks, trying to scan the whole address space including all our monitors. Note that some sub-scenarios could exist. For instance, someone scanning only CPS darknet ports and CPS honeypots or darknet generic (conventional) ports and CPS honeypots, etc. Identifying and distinguishing between scenarios could be tricky and require several activities. For instance, a random scanner could reach out to our CPS honeypot monitor indiscriminately. The only way to find whether the source is targeted (scenario 1) or random (scenario 2) is to check the activities that follow. For instance, a random activity could not reach out to a severe case (communication level as will be discussed in Section IV-C). Back to Figure 4, it means that 674 unique IP addresses were found in both datasets. This category of IP addresses is most probably under scenario 3. In our opinion, the most critical activities are within scenario 1; when a user tries to target only CPS honeypot or darknet CPS services.

B. CPS Abused Ports and Services

Our findings based on the CPS honeypot monitors show that IPMI is the most communicated service. In fact, more than half (51.38%) of our honeypot-based dataset includes IPMI. This phenomenon is reasonable as IPMI is a commonly used SCADA port is vulnerable due to its weak security of the Baseboard Management Controllers (BMCs). Note that IPMI service has approximately 40 Common Vulnerabilities and Exposures (CVE) since 2004 including 9 in 2019 as per the MITRE Corporation[1]. Apart from its vulnerability, this port is also a centralized system that is able to control subsystems. Therefore, an attacker having control of this service is capable of conducting significant control and damage

[1] https://cve.mitre.org/

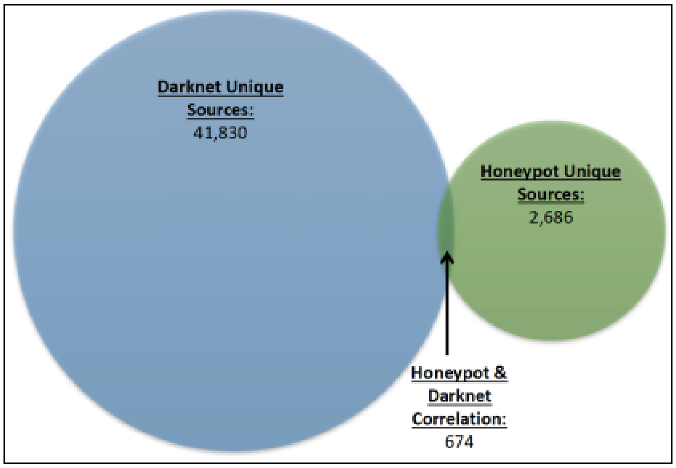

Fig. 4. Darknet and Honeypot Correlation

TABLE I
Top Targeted Ports/Services

Top Honeypot CPS Ports	Top Darknet CPS Ports	Top Darknet Conventional Ports
IPMI 623 (51.38 %)	IPMI 623 (56.29 %)	Others (64.13 %)
S7Comm 10201 (24.46 %)	BACnet 47808 (26.34 %)	Telnet 23 (7.01 %)
Modbus 502 (13.21 %)	Guardian 10001 (9.64 %)	Microsoft-DS 445 (5.45 %)
BACnet 47808 (5.33 %)	Kamstrup_mngt 50100 (6.18 %)	HTTP 8080 (5.29 %)
IEC104 2404 (4.07 %)	IEC104 2404 (0.4 %)	IPMI 623 (4.83%)
Kamstrup_ch_a 1025 (0.59 %)	Kamstrup_ch_a 1025 (0.34%)	Huawei 37215 (4.71 %)
Guardian 10001 (0.50%)	S7Comm 10201 (0.33%)	SMTP 25 (3.15%)
Kamstrup_ch_b 1026 (0.28%)	Kamstrup_ch_b 1026 (0.29%)	SSH 22 (2.76 %)
Kamstrup_mngt 50100 (0.18%)	Modbus 502 (0.18%)	HTTP 80 (0.96%)

over several machines such as a botnet. The second most targeted service is S7comm (S7 Communication), which is a commercial proprietary protocol for Siemens. This service runs between programmable logic controllers (PLCs) of the Siemens S7-300/400 family. S7comm occupied 24.46% of our honeypot data. The third targeted service is Modbus is a serial communications protocol originally published by Modicon (now Schneider Electric) in 1979. This service is for use with PLCs as well. Modbus has become a de facto standard communication protocol. This service is widely used today to connect industrial electronic equipment and it was found in 13.21% of our collected honeypot data.

As far as darknet data analysis, the results reveal that also IPMI is the most targeted CPS service with a distribution of 56.29% in our dataset. However, BACnet, which is a communication protocol for Building Automation and Control (BAC) networks have found second with a distribution of 26.34% in our dataset. BACnet has about 20 vulnerabilities since 2010 including 3 in 2020 against Cross-Site Scripting (XSS) and backdoor attacks. The third on the list is the Guardian AST service, which is a gas tank monitoring system. It consists of 9.64% of our darknet collected data. This service has been abused in the past several times targeting thousands of gas stations in the United States.

In regards to conventional (non-CPS related services), the most common activities have targeted Telnet service. Telnet is extremely vulnerable as it has no encryption and its packets are sent as plain text. The port is mainly used as a remote management connection and therefore an attacker gaining control over Telnet can cause severe damages. It is noteworthy to mention that the IPMI is the only CPS service that has been found among the top conventional darknet abused services. Table I provides a summary of the top targeted services per monitoring system.

C. Attack Severities

In order to label the severity of attacks we used a customized version of the techniques used in [13]. In a nutshell, session initiation is labeled as medium severity, requests such as HTTP GET are considered high severity, and finally traffic communication (data transfer) is labeled critical. Figure 5 depicts an overview based on the honeypot main investigated services. IPMI has been found critical in the majority of its communication, whereas BACnet is distributed almost equality between medium and high severity. The majority of S7Comm communication has been found high in terms of severity. In addition, Modbus traffic are distributed between high and critical, with majority to the latter. Our investigation reveals that IPMI has probably the most targeted attacks not just in terms of numbers, but also severity. Furthermore, the lack of critical severity label within BACnet and S7Comm is because these services run on top of UDP. As such, no connection setup and data exchange. Last but not least, Modbus has been found to be the only CPS service that has all variants of attack severity.

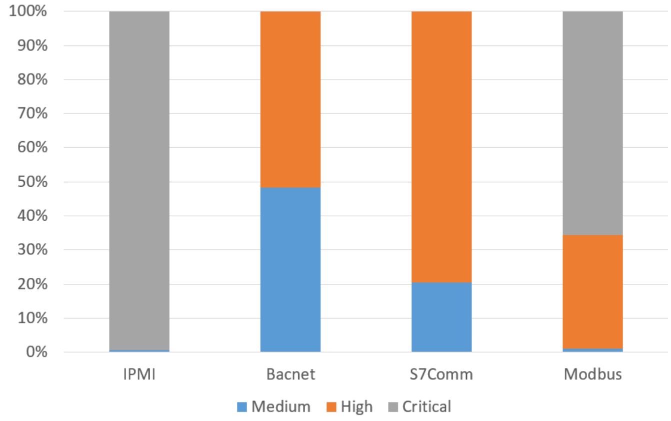

Fig. 5. Attack Severity in Honeypot Data

Please note that in this paper, we have elaborated on the severity as per the honeypot data only. We kept the severity of darknet data for future work.

D. Validation

Cross-examination with a trusted third-party dataset is required to validate our data. In addition to manual inspection, our findings are compared to results from a trusted third-party (AbuseIPDB)[2]. The latter provides a centralized blacklist and find IP addresses that have been involved with suspicious and malicious activity on the Internet.

The source IP addresses that originate network communication to our monitors are extracted separately from the Darknet and Honeypot systems. Subsequently, a script checks the occurrence and compares activities within our dataset and AbuseIPDB. We leverage the abuse confidence indicator from the latter service to validate our results. This confidence indicator is a very conservative value between 0 and 100, with 0 being not abusive and 100 as being extremely abusive. Optimizing threshold selection is kept for future work. However, in our experiments, we selected a conservative and balanced threshold of 50% to validate our findings. For example, once our system detects an IP address, we check the aforementioned trusted third party dataset, if the same IP address has more than 50% abuse confidence, then we label the IP address and its network flow as malicious. Taken into account the aforementioned thresholds, 43% of source IP addresses in our darknet data have been found and in the AbuseIPDB site with an abuse confidence of more than 50%. Furthermore, 25% of the source IP addresses in our darknet dataset have been found in the AbuseIPDB with at least 50% of abuse confidence levels. The remaining (unreported) IP addresses could fall into the following scenarios: 1) unreported to the trusted third-party; 2) undetected by other parties; 3) false positive, which means activities that we have labeled as a threat and are in fact not; or 4) found but have an abusive confidence value of less than 50%. Except for the latter scenario, identifying the correct reason may involve a further analysis, investigation, and collaboration with other parties. We left this task for future work.

It is noteworthy to mention that we have filtered out Internet-wide scanners to focus more on threat flows and reduce false positive. However, for measurements only, we list below the number of activities identified by well-known Internet-wide scanners such as Shodan, ShadowServer, IPIP.net, Arbor Observatory, and Censys.

TABLE II
HONEYPOT AND DARKNET HITS ON OUR MONITORS

Scanner	Packets on Darknet	Packets on Honeypot
Shodan	319203	1087632
ShadowServer	537921	51285
IPIP	402126	59049
Arbor Observatory	225813	24017
Censys	523045	34491

As shown in Table II, Shodan is the most active and focused Internet-wide scanners in terms of targeting CPS interactive services (e.g., many-to-one distributed port scan), whereas ShadowServer and Censys are the most active in generic network scans (e.g., one-to-many port sweep).

V. DISCUSSION AND LIMITATIONS

While the proposed system highlighted several contributions, there were some limitations to our approach. First, building cyber security capabilities by deploying monitors or sensors require few challenges in terms of collaboration between different entities, namely, Internet Service Providers, Governments, space providers and IT support where the monitors are deployed. In our case, our monitors are based at our academic institutions. It was challenging to convince the IT administration and leaders to approve our project. Furthermore, it was also challenging to access the critical infrastructure (e.g., data center), where we hosted our servers. Last but not least, getting approval/authorizations to deploy monitors and comply with the city and government policies was the most demanding task.

In addition, the dataset used is limited to the scope we cover and the duration (period) of our analysis. In other words, we can only detect threats that target our dataset and network infrastructure during the analyzed period; any traffic beyond our monitoring scope cannot be seen. Furthermore, our data is limited to a year and a half. As such, analyzing more comprehensive (greater than 2-year data) is left for future work.

Although our honeypot monitors mimic SCADA and industrial control systems, it is almost impossible to assess the real attack impact (damage) on physical infrastructure without having one. Therefore, our model helps in investigating CPS threats at the initial phases (reconnaissance, infiltration) but not the last ones (exploits and damage). Correlating our monitors with real testbeds are kept for future work. Several data is received from well-known organizations such as Shodan, ShadowServer, and research institutions. Identifying and filtering between benign and malicious data in systematic and automated manners are done manually. We are currently working on automating this process using advanced time-series techniques.

VI. CONCLUSION AND FUTURE WORK

Monitoring the cyberspace is a fundamental requirement to protect the Internet. There have been a few ways to protect CPS and OT systems offline or on-site protection but not enough from a cyber perspective. As such, we proposed a Big Data fusion model that leverages a correlation engine between two types of trap-based network monitoring systems, namely, honeypot and darknet. Furthermore, we have conducted a comparative study to evaluate the accuracy and performance of these two aforementioned systems. This contribution has had a number of promising findings that could aid in defense of the Cyberspace as well as cyber-physical systems. In fact, results from the honeypot deployments captured a relatively

[2] https://www.abuseipdb.com/

large amount of cyber activities originating from Europe. Such activities are dedicated to probing SCADA ports/services in a critical infrastructure environment. Honeypot data are found to be more stable than darknet. The latter though are more attractive monitors to hackers and easier to secure and hide. The IPMI is the number one targeted CPS service. It is in fact the only CPS service that has been found among the top conventional darknet abused services. As far as darknet deployments, Asian countries are the most active sources. Geolocating sources indicate research institutions such as Research Agency (IRA) involvement and well known Internet-wide scanners. Several discussions have been listed and our future work includes the analysis of a larger and more recent dataset and the integration of physical testbed into our monitoring and detection model.

VII. Acknowledgment

The authors would like to thank the Dubai Electronic Security Center (DESC) for funding this project and providing support and guidelines to achieve the objectives.

References

[1] Christian Czosseck, Rain Ottis, and Anna-Maria Talihärm, "Estonia after the 2007 cyber attacks: Legal, strategic and organisational changes in cyber security," *International Journal of Cyber Warfare and Terrorism (IJCWT)*, vol. 1, no. 1, pp. 24–34, 2011.

[2] Nicole Perlroth and Clifford Krauss, "A cyberattack in saudi arabia had a deadly goal. experts fear another try," *New York Times*, vol. 15, 2018.

[3] Claude Fachkha and Mourad Debbabi, "Darknet as a source of cyber intelligence: Survey, taxonomy, and characterization," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1197–1227, 2015.

[4] Sameera Almulla, Elias Bou-Harb, and Claude Fachkha, "Cyber security threats targeting cps systems: A novel approach using honeypots," in *The Twelfth International Conference on Emerging Security Information, Systems and Technologies, SECURWARE*. IARIA, 2018, pp. 85–91.

[5] Samuel Litchfield, David Formby, Jonathan Rogers, Sakis Meliopoulos, and Raheem Beyah, "Rethinking the honeypot for cyber-physical systems," *IEEE Internet Computing*, vol. 20, no. 5, pp. 9–17, 2016.

[6] Lukas Rist, Johnny Vestergaard, Daniel Haslinger, A Pasquale, and J Smith, "Conpot ics/scada honeypot," 2013.

[7] Claude Fachkha, Elias Bou-Harb, Anastasis Keliris, Nasir D Memon, and Mustaque Ahamad, "Internet-scale probing of cps: Inference, characterization and orchestration analysis.," in *NDSS*, 2017.

[8] Claude Fachkha, "Cyber threat investigation of scada modbus activities," in *2019 10th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*. IEEE, 2019, pp. 1–7.

[9] Giuseppe Bernieri, Mauro Conti, and Federica Pascucci, "Mimepot: a model-based honeypot for industrial control networks," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. IEEE, 2019, pp. 433–438.

[10] Michael Haney, "Leveraging cyber-physical system honeypots to enhance threat intelligence," in *International Conference on Critical Infrastructure Protection*. Springer, 2019, pp. 209–233.

[11] Samuel Lewis Litchfield, *HoneyPhy: A physics-aware CPS honeypot framework*, Ph.D. thesis, Georgia Institute of Technology, 2017.

[12] Claude Fachkha, Elias Bou-Harb, and Mourad Debbabi, "Fingerprinting internet dns amplification ddos activities," in *2014 6th International Conference on New Technologies, Mobility and Security (NTMS)*. IEEE, 2014, pp. 1–5.

[13] Charlie Scott and Richard Carbone, "Designing and implementing a honeypot for a scada network," *SANS Institute Reading Room*, p. 39, 2014.