THE GEORGE WASHINGTON UNIVERSITY
THE ELLIOTT SCHOOL
OF INTERNATIONAL AFFAIRS

THE INSTITUTE FOR MIDDLE EAST STUDIES

IMES CAPSTONE PAPER SERIES

Cyber Threat Indicators and the Middle East Cyber Attack Database

Gabrielle Barbour and Cory Stephens

April 2016

THE INSTITUTE FOR MIDDLE EAST STUDIES
THE ELLIOTT SCHOOL OF INTERNATIONAL AFFAIRS
THE GEORGE WASHINGTON UNIVERSITY

Table of Contents

ACKNOWLEDGMENTS **3**

ABBREVIATIONS **4**

EXECUTIVE SUMMARY **5**

PAPER OUTLINE 5

SECTION I – INTRODUCTION **8**

SECTION II - THE MIDDLE EAST CYBER ATTACK DATABASE **12**

A BRIEF HISTORICAL INTRODUCTION TO THE GLOBAL TERRORISM DATABASE 12

THE GTD AS A MODEL FOR THE MECADB 15

Limitations of Single Incident Determination *17*

SECTION III - DATABASE DESIGN, COLLECTION, AND TESTING **20**

Incident Table *20*

Incident Origin Table *21*

Incident Target Table *22*

Incident Perpetrator Table *24*

Incident Payload, Exploit, and Propagation Tables *25*

Incident Source Table *26*

DATA COLLECTION METHODOLOGY 27

SECTION IV - STATISTICAL ANALYSIS AND THE MECADB CYBER THREAT INDICES **30**

CYBER ATTACK SOPHISTICATION INDEX (SI) 30

CYBER ACTOR THREAT INDEX (TI) 31

MECADB QUANTITATIVE FINDINGS 32

MECADB ANALYTIC FINDINGS: THE DOMINANT GROUPS 33

CONCLUSION: CREATING A SUSTAINABLE GLOBAL CYBER ATTACKS DATABASE **37**

BIBLIOGRAPHY **39**

Acknowledgments

We would like to acknowledge the following individuals to whom we are indebted in the development and pursuit of our project: Michael Jensen, START/GTD; Ben Webster and Katie Bakewell, Data Scientists; Christopher Davenport and John Stephens, Software Engineers. Thank you for your support and guidance throughout every phase of the Middle East Cyber Attack Database. Additionally, we thank our families, specifically Charisse Stephens and Ryan Brennan, who have sacrificed alongside us, and without whom this research would not have been possible.

Abbreviations

CI	Source Confidence Index
DHS	Department of Homeland Security
DDoS	Distributed Denial of Service
DoD	Department of Defense
DoS	Denial of Service
GCAD	Global Cyber Attacks Database
GIS	Geographical Information Science/System
GTD	Global Terrorism Database
IAEA	International Atomic Energy Agency
ICT	Information Communication Technology
IS	Islamic State
MECAdb	Middle East Cyber Attack Database
MENA	Middle East/North Africa
OSINT	Open Source Intelligence
SEA	Syrian Electronic Army
SI	Cyber Attack Sophistication Index
START	Study of Terrorism and Responses to Terrorism
TI	Cyber Actor Threat Index
US-CERT	US Computer Emergency Readiness Team

Executive Summary

Cyber attacks originating from the Middle East and North Africa (MENA) region are a growing concern to corporate, political, and national security interests for countries within and outside the MENA region. An open-source database tracking cyber activities in the region will produce the raw data required to understand this relatively new phenomenon with quantitative analytic tools. The database developed and designed by the authors—with guidance from software engineers, data scientists, and data collection specialists—is a working prototype that, once fully implemented, will aid in describing, understanding, and reporting malicious cyber attacks and their perpetrators**.**

The Middle East Cyber Attack database's (MECAdb) design and initial implementation is the culmination of research by Cory Stephens and Gabrielle Barbour under the advisement of Dr. Michael Jensen, the data collection manger of the National Consortium for the study of Terrorism and Responses to Terrorism's (START) Global Terrorism Database (GTD). Considering the nature of our research interest in actors behind the attacks, the design and theoretical foundation of the MECAdb takes inspiration from the GTD. The prototype database is the culmination of our multidisciplinary study of the Middle East at the George Washington University Elliott School of International Affairs. As a proof of concept, the MECAdb demonstrates the viability of our analytic project and once fully implemented, it has the potential to become a key resource in understanding the impact of hackers in the Middle East.

Paper Outline

This project is divided into four main sections. First, an introduction explains the origins of the MECAdb. This section briefly describes the cyber landscape of hackers in the Middle East. We assess how the current scholarship on information and communications technology excludes consideration of hackers and address why we believe this analytical gap exists. We then discuss discuss what resources exist in order to aggregate data on cyber attacks in the region and how this information can be used to quantify the phenomenon.

The second section offers a brief historical introduction to the GTD and demonstrates how the GTD served as a model for the MECAdb. We discuss the history and development of the GTD over time and several of the analytic products that rely on information only found in and through the efforts of the GTD. We argue that given resources and time, the MECAdb can become the first comprehensive source of information on cyber attacks originating from the Middle East and that it has the same analytic potential as the GTD.

The third section outlines the database structure, including inclusionary criteria and data collection methodology. We approach the structure of the database through a set of *incident* tables which record incident metadata (dates, summary, inclusion criteria, and whether the attack is international or domestic), origins and targets (including geospatial information when possible), perpetrators, cyber weapons, and sources. Finally, we discuss the data collection methodology used in the prototype and areas for development.

The fourth and final section explores possible avenues of statistical and quantitative analysis of the MECAdb. We discuss the Cyber Attack Sophistication Index (SI) and the Cyber Actor Threat Index (TI), two analytic indices developed in collaboration with data scientists to understand how the MECAdb can be used to produce meaningful, original, and actionable analysis relevant to broader regional studies. We then include statistical findings, identified patterns, and geographical visualizations of the data contained in the prototype database. While these can only serve as examples of analysis, they highlight both the potential of a fully implemented database and necessary development to reach that potential.

Section I – Introduction

In the midst of the popular protests now known as the Arab Uprisings, a group of cyber mercenaries formed across the Middle East region. Comprising an estimated thirty native Arabic speaking individuals, the Desert Falcons represent the "first known Arab group to develop and run a full cyber-espionage operation."[1] At around the same time, another "community of cyber warriors" formed in Syria in support of Bashar al-Assad's political regime. The self-proclaimed Syrian Electronic Army (SEA) began their online presence with a series of high profile web defacements and propaganda campaigns targeting websites and social media accounts they considered critical of the Assad regime.[2] These and other hacker collectives have proliferated across the Arab world and cyber security professionals have taken note. On March 22, 2016, the United States District Court charged three members of the SEA with crimes ranging from conspiracy, to money laundering, violation of sanctions, and extortion. Two days later, seven Iranians were indicted for their participation in campaigns targeting the US financial sector.[3]

These three cases—The Desert Falcons, the SEA, and the indictments of individuals connected to SEA and the state of Iran—place the Middle East alongside sophisticated and increasing influential cyber threats. The rapid proliferation of information

[1] Kaspersky Labs. *The Desert Falcons Targeted Attacks.* (Kaspersky Labs, Moscow: Kaspersky Labs, 2015), 4.

[2] Ahmed Al-Rawi. "Cyber warriors in the Middle East: The case of the Syrian Electronic Army ." *Public Relations Review* 40 (April 2014).

[3] *United States of America v. Ahmad Fathi et al.* (United States District Court Southern District of New York, March 24, 2016).
United States of America v. Ahmad 'Umar Agha and Firas Dardar. 1:14-MJ-292 (United States District Court for the Eastern District of Virginia, June 12, 2016).
United States of America v. Peter Romar and Firas Dardar. 1:15-mj-00498-MSn (United States District Court for the Eastern District of Virginia, September 29, 2015).

communication technology (ICT) and its users in the greater Middle East (including Turkey and Iran) is not a new subject to regional analysts and academics. Political scientists such as Steven Heydemann and Marc Lynch have discussed how Middle East authoritarians use the Internet and ICT to "upgrade" their regimes[4] and how activists use the same tools to organize and express discontent.[5] Further, Yeslam Al-Saggaf at Charles Stuart University draws upon the work of several social scientists in his ethnographic work to understand how online news sites contribute to an Arab online public sphere, concluding that reader's use of online message boards hosted by two news websites "transforms them from a passive audience to authors of media content" and that these sites can therefore be considered—with some reservations—part of an online Habermassian public sphere.[6] The study of the Internet in the Arab world is a multidisciplinary area of interest to diverse fields of research.

It is apparent that analysts are concerned with how the Internet is being used in the Middle East. It is also evident that hackers and other so-called "bad actors" are proliferating at a rapid pace in the last decade. Yet, when one seeks in-depth, analytical consideration of how and why groups like the SEA and the Desert Falcons use and misuse cyber infrastructure, the literature is sparse; and that literature which does exist is often overshadowed by industry reports published by cyber security firms such as Kaspersky Lab, Cylance, and Mandiant FireEye. In an informal conversation about the

[4] Steven Heydemann. *Upgrading Authoritarianism in the Arab World.* Analysis Paper (The Saban Center for Middle East Policy, The Brookings Institution, The Brookings Institution, 2007), 18-23.

[5] Marc Lynch. "Media, Old and New." In *The Arab Uprisings Explained: New Contentious Politics in the Middle East*, edited by Marc Lynch (New York, New York: Columbia University Press, 2014), 93-109.

[6] Yeslam Al-Saggaf. "The Online Public Sphere in the Arab World: The War in Iraq on the Al Arabiya Website." *Journal of Computer-Mediated Communication* (International Communication Association) 12 (2006). 311-334.

2014 high-profile hack of Sony Pictures Entertainment, a former security specialist with FireEye related that such reports were always reviewed by the firm's advertising arm to leverage the work's potential to attract new clientele. Despite this bias, however, these reports are one of very few sources for in-depth, publicly available information on cyber threat actors, and while they often have secondary motives, the information is valuable to understand regional actors.

The set of three indictments published by the United States District Court represent an alternative source of information on online bad actors from the Middle East - government intelligence and law enforcement agencies. In the cases cited above, the Federal Bureau of Investigation (FBI) documents dozens of attacks attributed to the named defendants. The affidavits supporting the charges include target information, malware and techniques used, as well as motives. These indictments are, likewise, based on presumably reliable data. They nevertheless have the bias of establishing probable cause for criminal activity and therefore are limited in their analytic value.

These two examples represent, we believe, the difficulty of studying hackers and hacker collectives in the Middle East. The most comprehensive and reliable data sources are either private, proprietary, or classified. Additionally, because of the institutional priorities of the owners of the data, certain types of analysis are favored over others. In the case of private security firms the institutional mandate is to protect client systems. This by default prioritizes system vulnerability and attack vector analysis over larger questions of attribution, motive, or broad socio-political impact. In the case of the intelligence community—who very likely *are* concerned with actors for national security and who have access to vast amounts of data—the weakness is a lack of robust systems for information sharing and granting technical information access to multidisciplinary analysts. While one could conceivably create contexts in which civilian analysts are granted access to classified intelligence for purposes of national security, it is considerably more difficult to envision the same for non-security contexts such as anthropology, ethnography, or other social sciences.

Studying hackers and hacker collectives in the Middle East is a new frontier in regional studies and has proven to be a difficult field to enter. However, even if the most in-depth data related to cyber attacks is locked behind intellectual property rights and government classification, we argue that sufficient publicly available information exists for academics and independent analysts to produce meaningful and rigorous conclusions. In the age of big data and statistical analysis, the barrier to entry is not a lack of information. It is a lack of the proper tools to access the information in efficient and meaningful ways. In fact, given analytic tools and a reliable, centralized repository of public information, we believe that new, multidisciplinary studies will emerge to fully integrate these important actors into their broader contexts. To this end, we have designed and created the first open-source relational database of cyber attacks originating from the Middle East and North Africa region.

Section II - The Middle East Cyber Attack Database

The MECAdb itself and the structural procedures surrounding it collect and analyze sources describing cyber attacks originating in the MENA region. It likewise parses out individual data points and relates those points to one another in meaningful ways. Its primary purpose and greatest potential is to understand the broad impact and development of Middle East hacking collectives. Once fully implemented, the MECAdb will provide the raw data required to answer questions such as:

- What is the most frequent type of attack originating from a specific country?
- What factors predict whether a group will become more or less sophisticated?
- How do geopolitical events affect the number, nature, and/or sophistication of attacks?

Because our primary goal is to understand the actors, the geopolitical impact, and the role of cyber attacks in international security, the MECAdb is fashioned after a similar project: The Global Terrorism Database (GTD), maintained by the University of Maryland's National Consortium for the Study of Terrorism and Responses to Terrorism (START).

A Brief Historical Introduction to the Global Terrorism Database

In their introduction of the GTD, Gary LaFree and Laura Dugan note that the GTD "contributes to [open-source databases on terrorist events] by providing for the first time a comprehensive collection of terrorist events including both domestic and international incidents for several decades."[7] From the time of its introduction, the database has grown from approximately 70,000 events to the 140,000 it contains now. Relying only on open

[7] Gary Lafree and Laura Dugan. "Introducing the Global Terrorism Database." (*Terrorism and Political Violence* 19, no. 2 July 2007), 198.

sources, the database collects information on event dates, geographical locations, weapons, and the nature of targeted entities. Additionally, it aggregates data on victims of terrorist attacks and attributive information when possible. According to the database website, the GTD is "currently the most comprehensive unclassified data base on terrorist events in the world."[8] Citing over 4 million articles from over 25,000 news sources, the project participates in the "effort to increase understanding of terrorist violence do that it can be more readily studied and defeated."

Lafree and Dugan note that the database has been useful in policy impact assessments, future risk of terrorist activity, and has been used to recommend anti-hijacking policy, criminal justice policy, and impact analysis of specific attacks.[9] Generally speaking, there are three types of users who access the contents of the GTD: the general public, journalists and other analysts, and counter-terror experts (including government and military personnel). These three types of users can represent the depth and level of sophistication for which the GTD can be used. For the general public it allows them an easily accessible and understandable means by which to educate themselves. When the general public has access to quality information and analysis they make informed voting and advocacy decisions, leading to sound public policy. For journalists and other analysts who already have knowledge on terrorist activities, it allows them to find trends, draw connections and conclusions, and propose policy recommendations. The Institute for Economics and Peace's Global Terrorism Index (GTI), which "provides a comprehensive summary of the key global trends and patterns in terrorism," is a tangible example of

[8] National Consortium for the Study of Terrorism and Responses to Terrorism (START). "Overview of the GTD". *Global Terrorism Database.* (June 1, 2015).

[9] Lafree and Dugan, 198.

analysis in this category.[10] In many cases, researchers have created specialized datasets to analyze specific phenomenon. For example, a team of analysts at START published the *Profiles of Individual Radicalization in the United States* (PIRUS), which categorizes more than 1,500 individuals as Islamist, far right, or far left extremists. Another group, the Project on Violent Conflict, out of the Rockefeller College of Public Affairs & Policy at the University at Albany created the Big, Allied, and Dangerous (BAAD) list which identifies groups and their ideological and state affiliations.

At the highest levels of analysis, experts in terrorism and responses to terrorism—whether law enforcement, the intelligence community, or the military—utilize the raw data in the GTD and subsequent analyses to make tactical decisions, define priorities, and shape their missions. The most tangible government product that uses GTD data is the *Country Reports on Terrorism,* published annually by the State Department, which has direct impacts on funding decisions.

As the GTD continues to grow, it reaffirms itself as an authoritative source for quantitative analysis of terrorist attacks. It has been used in some cases to challenge and in others confirm claims made regarding terrorist activity and it allows researchers identify, compare, and contrast trends in terrorist activity. These trends are critical at all levels of analysis including everyday journalism in the world's most reputable newspapers, including the *New York Times, The Guardian,* and *The Washington Post.* We

[10] Institute for Economics & Peace. *Global Terrorism Index 2015: Measuring and Understanding the impact of Terrorism.* (Institute for Economics & Peace, http://economicsandpeace.org/wp-content/uploads/2015/11/Global-Terrorism-Index-2015.pdf, 2015), 2.

envision the MECAdb to follow the trajectory of this path and be likewise utilized. The true value of the both the GTD and the MECAdb lies in their ability to transform tactical, short-term responses into strategic longer-term pursuits at understanding a group or situation at large.

The GTD as a Model for the MECAdb

The MECAdb aspires to be the first comprehensive, public source of information on cyber attacks in the Middle East. We aim to develop our database to a point where the same scholars who use the GTD to study terrorists can use the MECAdb to study hackers. Ultimately, the MECAdb tracks a different phenomenon than the GTD. Nevertheless, the basic unit of analysis and the research goals share enough characteristics to justify using the GTD as a starting point. Both hackers and terrorists have ideological frameworks, political affiliations, and pursue practical goals. Both databases seek to collect and organize open source information in order to describe events in detail, but also relate individual pieces of information to one another in meaningful—and new—ways. We do not equate hackers and terrorists—or even comment on so-called "cyberterrorists"—but we do seek to leverage technology used to study terrorists to understand the relatively new world of Middle East hackers. We acknowledge the etymological weight, history, and associations implicit in terms such as "attack," "threat", and "weapon" which here to refer to inherently non-violent tools and actions. While an in-depth discussion about the rhetorical nature of our exact terms is outside of the scope of this paper, we understand these cannot be used in isolation of their connotations and their mere presence in the project affects the tenor of a larger debate of how we discuss cyber actions. Nevertheless, the terms we use below are widely used in the cyber security industry—whether private, military, or government—and we believe they are not out of place in the present context.

Before describing the how the database attempts to meet these goals, we must consider the exact definitions and the nature of the sources we intend to use. The MECAdb defines a cyber attack "as any deliberate, illegal use of a cyber weapon against a target network with the intention of causing unauthorized physical or digital effects." Where a cyber weapon is defined as "a set of electronically or physically delivered code that includes one or more propagation methods, exploits, and payloads."[11] We explore the implications and reasoning for these definitions below in context of how they are used in the database.

Both the GTD and the MECAdb rely on open-source data available to the general public. We understand, as do the designers of the GTD, that other sources of information exist and indeed may be more complete and reliable. Nevertheless, the GTD and intelligence services have successfully incorporated open-source intelligence (OSINT) in their analysis and we believe this will be the case once the MECAdb is fully developed.

The primary sources of information on cyber attacks are mass media reports, social media, government documents, self-reporting, and cyber firm security industry white papers. The diversity of sources provides a range of information and requires a significant amount of front-end analysis. In reviewing hundreds of reports during the design process, we believe the structure below is sufficient to accurately incorporate pertinent

[11] Cory Stephens. "Middle East Cyber Attack Database: Theoretical Foundations and Methodological Considerations." (*Unpublished.* Washington, D.C.: Unpublished, December 7, 2015), 12.

information. Nevertheless, throughout the development process we anticipate discovering limitations of the initial design and adjustments will be made as needed.

Limitations of Single Incident Determination

The MECAdb and the GTD share the basic unit of what is a "single" attack. In its codebook, the GTD notes the "incidents occurring in both the same geographic and temporal point will be regarded as a single incident, but if either the time of occurrence of incidents or their locations are discontinuous, the events will be regarded as separate incidents." [12] Intuitively, cyber attacks share the basic temporal and geographic characteristics as kinetic attacks. Indeed, the execution of malicious code occurs at a specific time and the physical systems processing the command (origin) and the execution (target) both occupy physical space. The nature of modern ICT infrastructure, however, complicates this determination and calls into question the value of thinking about cyber attacks in terms of traditional geography and time. For example, the SEA gained prominence by hacking into Twitter accounts of high profile news organizations, such as *the New York Times*, *The Washington Post*, and most notably, *The Associated Press*. However, these cyber attacks do not have geographical locations in the traditional sense since the hacked system (Twitter) differs from the attack target. In cases like Iran's alleged cyber attack that destroyed 30,000 computers at Saudi ARAMCO, the physical geography is more clear.

The variability of clearly defined borders is a fundamental aspect of cyber activity. In his

[12] START. *Global Terrorism Database Codebook: Inclusion Criteria and Variables.* Codebook, Global Terrorism Database, National Consortium for the Study of Terrorism and Responses to Terrorism, College Park: National Consortium for the Study of Terrorism and Responses to Terrorism, 2015.

work on human-computer-reality-interaction Michael Goodchild, a prominent geographer who has used computer mapping and Geographical Information Science (GIS), asserts that cyber-geographies are a "second age of geographical exploration."[13] He discusses how cyber-mapping is multidimensional and more complex than is traditional cartography and charting of physical space. [14] Goodchild's "multidimensional" perspective is useful in conceptualizing how we can consider multiple geographical locations of cyber activities depending on the scale of analysis. For our purposes, unless the source material specifically notes the effect of an attack on physical systems, the geographic location of the target headquarters is recorded as the location. If sources note physical systems or infrastructure, the location of the systems takes priority. With these considerations, the most precise measure of the number of individual attacks in the prototype MECAdb is the number of targets affected.

Limitations of Open Source Data Aggregation

Since the data contained in both the GTD and the MECAdb is "culled from news sources," the databases have a natural bias favoring newsworthy attacks.[15]Any conclusions drawn from the MECAdb must therefore consider the weaknesses of open-source intelligence and it should note the biases of the reporting entity whether financial—as is the case with corporate white papers—or sensational as is often the case with popular news and social media. The GTD devotes considerable resources to analyzing source reliability, and

[13] Michael F. Goodchild. "Rediscovering the World Through GIS." (National Center for Geographic Information and Analysis, 1998), 1.

[14] M.F Goodchild. "Geographic information Systems: today and tomorrow." (*Annals of GIS* 15, no. 1. 2009), 1.

[15] Gary LaFree et. al. *Building a Global Terrorism Database.* Grant Report (University of Maryland, U.S. Department of Justice, unpublished, 2006), 24.

while the MECAdb does not currently implement source reliability measures, it will be implemented as the database is developed.

Section III - Database design, Collection, and Testing

In the following section, we will briefly tour the design of the MECAdb version 0.1.

Database Structure

In order to be analytically useful and as descriptive as possible, the MECAdb relies on three types of tables: definitions, relations, and incidents. For the sake of brevity, we focus here on the set of nine incident tables used to describe individual incidents. We will address specific design challenges and decisions in the descriptions of each.[16]

Incident Table

The *incident* field contains metadata surrounding an individual attack. It includes fields that assign permanent incident identification numbers, initiation/resolution dates, attack preparation time, incident summaries, inclusion eligibility, related incidents, and domestic/international classification.

In order to be included in the MECAdb, an event must meet three inclusionary criteria:

1) The incident must have a political, social, economic, or religious goal.
2) The incident must successfully execute at least one payload.
3) The incident must be directly related to the Middle East in its origin, target, or perpetrator. The MECAdb includes the following greater Middle East countries: Algeria, Bahrain, Egypt, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Palestine, Qatar, Saudi Arabia, Syria, Tunisia, Turkey, United Arab Emirates, and Yemen.

The MECAdb *does not* attempt to capture every cyber attack in the region and it relies on the analysis of attacks by the entity reporting. Therefore, if the source material notes political or social institutions, reports religious motivations, or if the payload is designed

[16] An earlier version of this paper included specific technical information on the exact design and structure of the prototype database. These have been removed to protect the intellectual property rights of their creator.

to accomplish a similar goal, we determine the incident meets criteria 1. Through future development, the MECAdb will be able to make more precise determinations of intent. We do not include cyber crime intended only for personal gain. Similar to the GTD excluding other forms of violence, our primary goal is to consider the broader socio-political impact of hacker groups.

In some cases an individual event can be deemed analytically relevant to the goals of the database without meeting all three criteria. These attacks may be included in the database however, in such cases, the variable *doubt_inclusion* would indicate the event did not meet all three criteria.

A specific challenge we faced was how to include reported numbers of additional attacks with little or no information. For example, if a news report mentions the same actor hacked forty other targets in relation to the one reported, this is analytically useful but without more information, we were unable to include this data as separate incidents. Therefore, we created the *attack_multiplier* field to assign additional "weight" to the incident.

Incident Origin Table

The *incident_origin* table records reported countries, regions, and cities from which the attacks originated. Geospatial information is included to facilitate exportation of data into geospatial information system (GIS) databases. The MECAdb was designed to use the

GTD-created list of country codes designed by Gleditsch and Ward.[17]

Incident Target Table

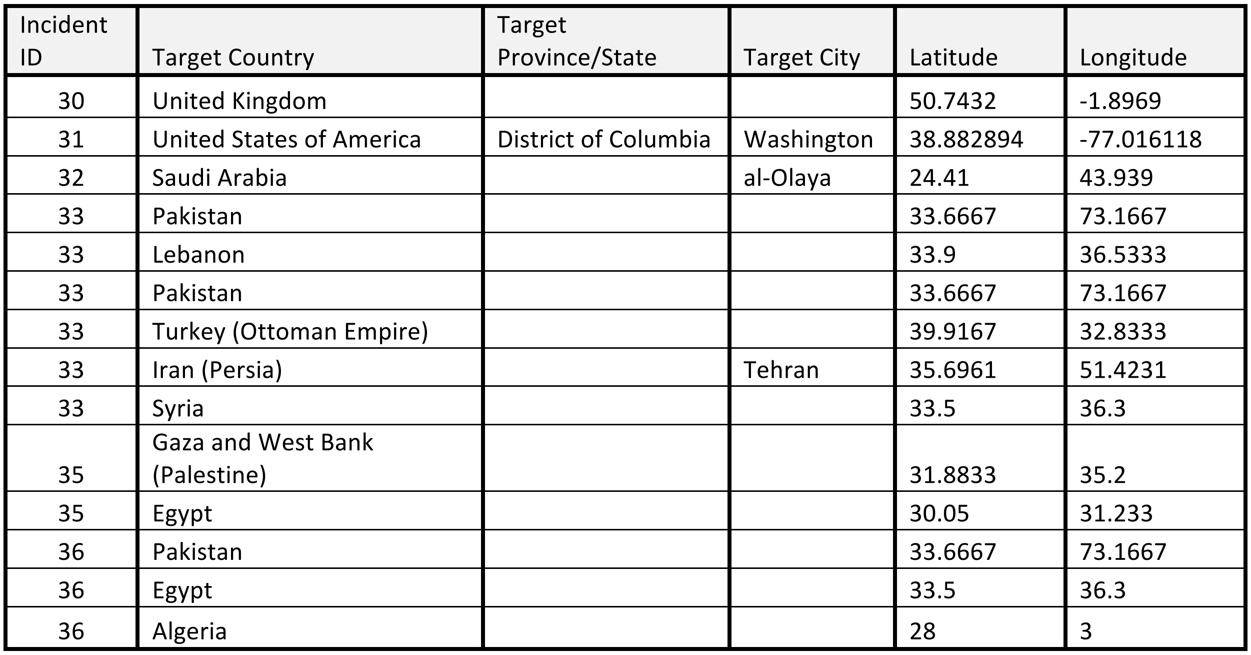
The *incident_target* table captures information on the targeted entity. Variables include: target country, province, and city with geospatial information when available; the target type and subtype; the corporate or individual entity; target nationality; the affected diameter in the case of attacks on infrastructure; as well as information on resulting deaths or damages.

The target information table is critical to our ultimate analytic goals. The level of detail included here will aid in making analytical judgments on the nature of the attackers and will help expose long-term political motivations.

We note here a key difference between how the GTD and MECAdb record data. In many cases, individual terrorist attacks target multiple entities. For example, in the case of the Paris attacks in November 2015, terrorists attacked six separate entities. Analysts must make the determination of whether these six attacks constitute a single attack or a series of attacks. In the event they are classified as a single incident, multiple targets must be recorded. The GTD solves this problem by creating an individual variable for each target, up to three targets. This is sufficient in the vast majority of cases but it poses a problem in the case of cyber attacks, which may target dozens or hundreds of entities in a single attack.

[17] Kristian S. Gleditsch and Michael D. Ward. "Interstate System Membership: A Revised List of the Independent States Since 1816." *International Interactions* 25 (1999): 393-413.

The problem of multiple values for a single variable is not unique to attack targets. Here, we encountered the issue of cyber attacks that may target dozens or hundreds of entities in a single attack. In an individual cyber attack, multiple entities can be targeted; it may include multiple origins, sophisticated multi-stage cyber weapons, and/or be reported by multiple sources. Rather than replicate the GTD solution of adding multiple variables (i.e. target1, target2, target3), which only solve the problem up to a certain number, we chose to leverage the power of relational data structures. With an entire separate table dedicated to *incident_target*, we can include unlimited targets, tied together using the incident id number. In *figure 1*, where *incident_id* = 30, a single target is indicated in the United Kingdom. Where *incident_id* = 33, a set of six targets are indicated each with their own country, target type, victim information, etc. When an analyst queries incident 33, she will be presented with a table indicating every target, every origin, every weapon, and every source cited.



Incident ID	Target Country	Target Province/State	Target City	Latitude	Longitude
30	United Kingdom			50.7432	-1.8969
31	United States of America	District of Columbia	Washington	38.882894	-77.016118
32	Saudi Arabia		al-Olaya	24.41	43.939
33	Pakistan			33.6667	73.1667
33	Lebanon			33.9	36.5333
33	Pakistan			33.6667	73.1667
33	Turkey (Ottoman Empire)			39.9167	32.8333
33	Iran (Persia)		Tehran	35.6961	51.4231
33	Syria			33.5	36.3
35	Gaza and West Bank (Palestine)			31.8833	35.2
35	Egypt			30.05	31.233
36	Pakistan			33.6667	73.1667
36	Egypt			33.5	36.3
36	Algeria			28	3

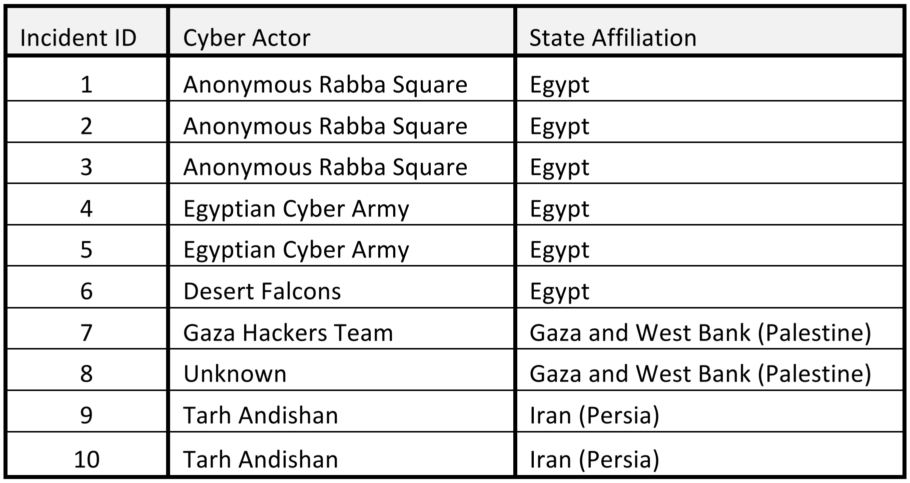
Figure 1: Example of how unlimited numbers of incident targets can be included in the MECAdb.

The current version of the MECAdb's *incident_target* table utilized the GTD definitions of target types and subtypes. Throughout the course of the initial round of data entry, we found several critical flaws in our input design. Initially we assumed that target types could be derived from an incident subtype. Generally, we found that in many cases cyber attacks target entities that are not targeted by traditional terror attacks. Therefore we must revise the type/subtype definitions in future versions.

Incident Perpetrator Table

The *incident_perpetrator* table records information on the entity/entities that executed an attack. The database includes a definitional table called *cyber_actor* that provides a standardized list of actors that grows as new incidents are added. Additionally, the *incident_perpetrator* table indicates whether the cyber actor group name is self-given or whether it is assigned by the reporting entity, as is often the case when firms reveal anonymous groups. Indicators also include whether analysts assign attributive claims or the groups themselves claim responsibility for the attack. The table likewise addresses competing claims, motives, ransom paid/demanded, and the text of any claims.

State_affiliation and *state_actor* fields indicate reported information only. By only recording reported affiliations, analysts may query the data for information on a single group and receive information on every attribution reported. They may then form their own analysis about which competing claims are reliable based on external information or statistical analysis.



Incident ID	Cyber Actor	State Affiliation
1	Anonymous Rabba Square	Egypt
2	Anonymous Rabba Square	Egypt
3	Anonymous Rabba Square	Egypt
4	Egyptian Cyber Army	Egypt
5	Egyptian Cyber Army	Egypt
6	Desert Falcons	Egypt
7	Gaza Hackers Team	Gaza and West Bank (Palestine)
8	Unknown	Gaza and West Bank (Palestine)
9	Tarh Andishan	Iran (Persia)
10	Tarh Andishan	Iran (Persia)

Figure 2: Query output detailing the cyber actor ID, group name, and reported state affiliation

Incident Payload, Exploit, and Propagation Tables

The *incident_payload, incident_exploit*, and *incident_propagation* tables record information related to the cyber "weapons" used in the incident. As stated above, a cyber weapon for our purposes is a set of electronically or physically delivered code that includes one or more propagation methods (Pr), exploits (E), or payloads (P). This definition relies on the PrEP Framework developed by Trey Herr in his report *PrEP: A Framework for Malware and Cyber Weapons.* We incorporate his framework for two reasons. First, it is not overly technical, and second, it breaks apart a complex subject into specific, describable pieces. He says:

> The PrEP framework focuses not on process, but rather on characteristics of the tools being used—suggesting that all malware can be conceptualized as the combination of three components: a propagation method, exploits, and a payload…This modular approach breaks up the current dominance of vague and

> variously defined terms such as worm, trojan, and virus, to focus on the intrinsic characteristics of three functional elements which make up all malware.[18]

This is essential to the integrity of the database as the sophistication of an individual attack largely relies on the weapon used. An attack propagated through mass spamming of a phishing email - which includes an open-source, widely distributed exploit method and payload - is less sophisticated. A spear-phishing email directed at a single user accessed through social engineering that uses zero-day exploits to deliver multiple payloads that in turn open secondary propagation methods, exploits and payloads, is more advanced. Moreover, a payload that aims to replace a website's homepage with a simple text or picture message is not going to warrant the same sophistication ranking as a prolonged access disruption payload.

As with the *incident_target* table, each of the cyber weapon tables has the ability to record multiple malwares, exploits, and payloads. In addition to defining the nature of the weapons used, these tables record whether specific tools used zero-day exploits, the source of the exploits, and whether the tools used were publicly available or custom designed/programmed. These details will be critical in determining an incident's level of sophistication.

Incident Source Table

The *incident_source table* allows individual entries for each source reporting an incident. Each entry includes information on the type of source (newspaper, social media, blog,

[18] Trey Herr. *PrEP: A Framework for Malware and Cyber Weapons.* Thesis, Cyber (Security and Policy Research Institute , George Washington University, Washington, DC: Cyber Security and Policy Research Institute, 2014), 4.

white paper, government report, etc), whether the source was in Arabic, and citation information. As the current version of the database is in its prototype phase, no source reliability information is recorded. Implementation of source reliability indices will take place as the database is developed.

Data Collection Methodology

As stated, The MECAdb relies on publicly available, unclassified source materials, consisting of news articles, broadcast news, security company white papers, and social media. For tracking individual attacks, online searches of news articles across multiple languages and mediums were utilized. Instruments such as Google, Facebook, and Twitter were used to search for attack information. The most popular method for documenting attacks was individual news articles. These included newspapers across the Middle East and North Africa – written in Arabic, English, French, Turkish, Persian, and Hebrew - as well as Europe, Russia, and the Unites States – written in English, French, Spanish, Russian, and Italian.

Within these news articles, certain sources were more reliable than others. While this was not quantified in the prototype database, analysts decided which reports were deemed reliable at the time of input. In general, articles written by third parties – i.e., when a British newspaper reports on an attack done by Turkish actors against Israel – were considered more trustworthy than an article written by a news source located within either the target or origin country. The reasoning for this is rather intuitive; a disengaged third party has less motivation to exaggerate information. Likewise, certain print and electronic news sources are more reputable than others. Factors considered when deciding how trustworthy a news source include both qualitative and quantitative factors. Logistics such as readership, scope of coverage by the news source, and whether or not it is located in a country with heavy state censorship are some of the measurable aspects. Other less tangible variables include any obvious pro-/anti-government sentiments either for or against the host country and others associated with the attack in question. When entering incident information, if we felt that a bias in reporting was present, we identified the attack as questionable by marking the *doubt-inclusion* indicator.

Secondary to news sources are attacks whose coverage never make it to official news sites but are often claimed on social media accounts such as Facebook or Twitter. When an attack is claimed on either of these mediums and it neither provides evidence for the attack (i.e., a link to pastebin, a picture of the defacement on the claimed website, etc.) nor is it corroborated by a separate news source, then the attack was not deemed trustworthy enough to be included in the prototype MECAdb database. If an attack claimed on social media *does* provide proof, or is corroborated by a news source that references the social media claim, the attack will be included in the database with a note that it was not independently verified.

A third form of data collection utilized was white papers published by security firms. These papers proved extremely fruitful in providing in-depth analysis on groups, particularly the type of exploits used, attack vectors, and payloads. Individual white papers were entered into the database using single incident identification numbers with details recorded in multiple entries in the incident target and cyber weapon tables. With all of the above mediums for data collection, there were, infrequently, conflicting claims of group attribution. In some of these instances, it was simply varying names of the same group; in others they were entirely different groups. In the latter case, the MECAdb documented the group that was attributed to the attack at a more current date by the news source deemed most trustworthy.

Section IV - Statistical Analysis and the MECAdb Cyber Threat Indices

The main question that the MECAdb seeks to answer is whether an open source database of cyber attacks is a viable analytic tool for assessing the impact of individual actors on the region. In working toward accomplishing this, we collaborated with data scientists to explore two analytic indices. Each is designed to evaluate weights assigned to particular data fields, and through a series of algorithms, create quantitative measures of the relative sophistication of individual incidents and the threat levels of groups based on their aggregate activities. At this point in development, the indices are best viewed as proofs of concept and are not intended to represent the current cyber landscape in the Middle East. In future developments, the MECAdb will calibrate the analytic algorithms by consulting with cyber security industry experts.

Cyber Attack Sophistication Index (SI)

The SI considers the cyber weapons used—recorded as incident_propagation, incident_exploit, incident_payload—the nature of the malware, and the overall impact of the attack. The database includes detailed information on specialized tools, payloads, propagation methods, in addition to information on the final impact. This raw data can be leveraged by statistical analysis and data scientists to rank the severity of individual attacks.[19]

Rank	Incident ID	SI Score
1	87	0.7283
2	59	0.7250

[19] In an earlier version of this paper, we detailed how the SI here and TI below were calculated in this prototype database through example algorithms donated by our partner data scientists. These have been removed in the final version to protect the intellectual property rights of their creator.

3	34	0.7008
4	80	0.6900
5	6	0.6500
6	55	0.6400
7	132	0.6400
8	31	0.6200
9	13	0.6000
10	33	0.5900
11	65	0.5900
12	36	0.5900

Figure 3: Top Incidents Ranked by Sophistication index.

Cyber Actor Threat Index (TI)

The Cyber Threat Index is a measure of the overall sophistication level of a cyber actor. Using a dataset of all Sophistication Index scores attributed to a single cyber actor, the TI calculates a relative threat level in a specific moment in time. Using an example analytic algorithm, the MECAdb produced an example top ten list of threat actors.

Rank	Group Name	Avg SI Score	Max SI	# of Incidents	Cyber Actor TI
1	Syrian Electronic Army	0.147	0.690	15	0.612
2	Tarh Andishan	0.127	0.700	8	0.609
3	Turkish Ajan Group	0.130	0.640	5	0.590
4	Moroccan Ghosts	0.105	0.640	7	0.581
5	AnonGhost	0.161	0.490	5	0.550
6	Ankincilar	0.152	0.490	9	0.547
7	Morocco Agent Secret	0.152	0.490	5	0.547
8	Ayyildiz Tim	0.130	0.490	6	0.540
9	Human Mind Cracker	0.132	0.490	7	0.530
10	Izz ad-din al-Qassam Cyber Fighters	0.094	0.490	6	0.528

Figure 4: Top Ten Groups ranked by Cyber Actor Threat Index.

With regular snapshots of an actor's TI over time, statistical analysis may be used to measure the progression or digression of a group over its lifetime.

MECAdb Quantitative Findings

The initial findings discussed below are the direct results as reported by the prototype MECAdb. They reflect the findings as dictated by the present data and indices and are therefore meaningful when considering the viability of the MECAdb and highlighting areas for further development.

The database recorded both raw statistics as well as analytic deductions. First, it recorded 42 groups and 409 targeted attacks. Of these, 87% were access attacks and 16% were an escalation of privilege, demonstrating that the majority of attacks were of a lower sophistication level.

Quantifiably, the database counted the number of attacks originating from each nation.

Country	Number of Targeted Entities
Turkey	48
Iran	26
Unknown	21
Syria	15
Morocco	12
Gaza and West Bank	8
Tunisia	7
Egypt	6
Israel	6
Saudi Arabia	5
Kuwait	4

Lebanon	3
Canada	2
Netherlands	1
Libya	1

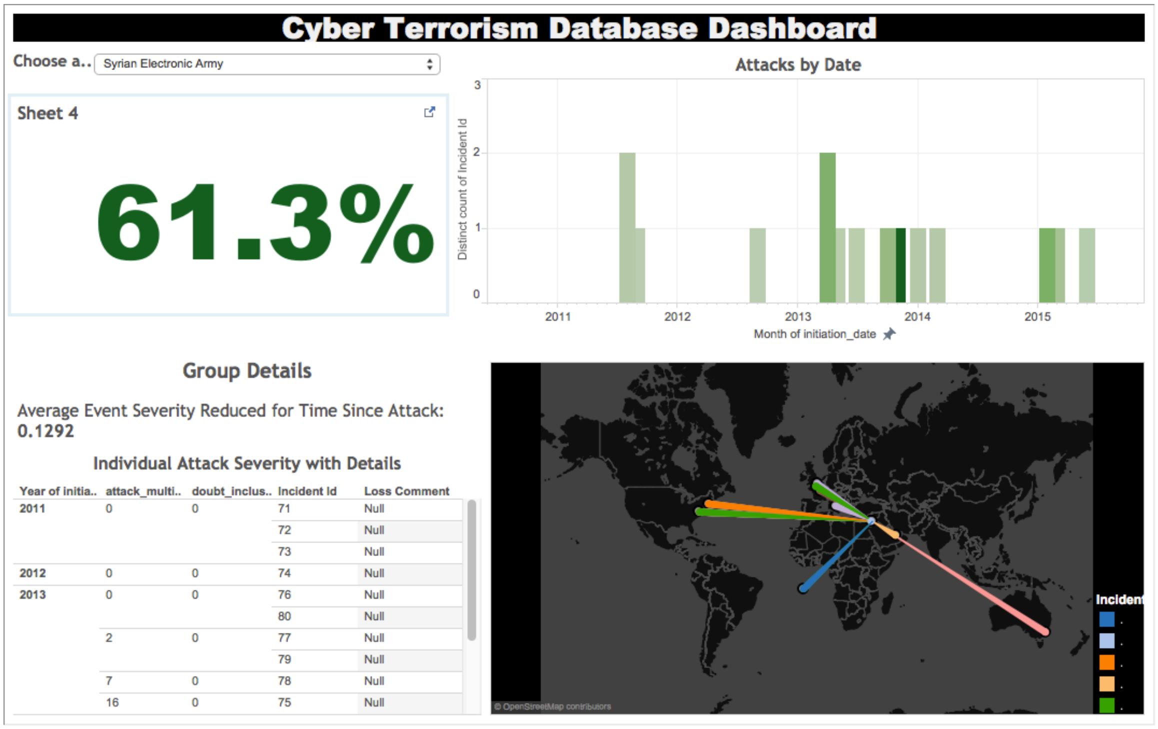
Figure 5: Number of Targets by Country

The database not only determined that the greatest number of attacks had an origin location in Turkey, but that also that attacks originating from Turkey were 82% more likely to have a Middle Eastern target versus all other nations as well as were 7.69 times more likely to be a domestic attack.

MECAdb Analytic Findings: The Dominant Groups

In addition to raw numbers, the database used a ranking algorithm to come to analytical conclusions. As indicated by the MECAdb, the SEA proved to be the most dominant actor, followed by Tarh Andishan, and the Turkish Ajan Group (See Figure 4 above.)

Here we find a discrepancy between the analytical findings of the authors and that of the MECAdb. By fleshing out how the database came to these conclusions, we find weaknesses in the current database that must be addressed in future versions. The data below shows the SEA, a group with an affiliation to the Syrian state, as a greater threat than Tarh Andishan, a group that is agreed upon by cyber experts to represent the Iranian government. However, as analysts with intimate knowledge of the attacks, having gathered them, we posit the opposite be true in regards to the number one and number two ranked spots.



Cyber Terrorism Database Dashboard

Choose a.. Syrian Electronic Army

Sheet 4

61.3%

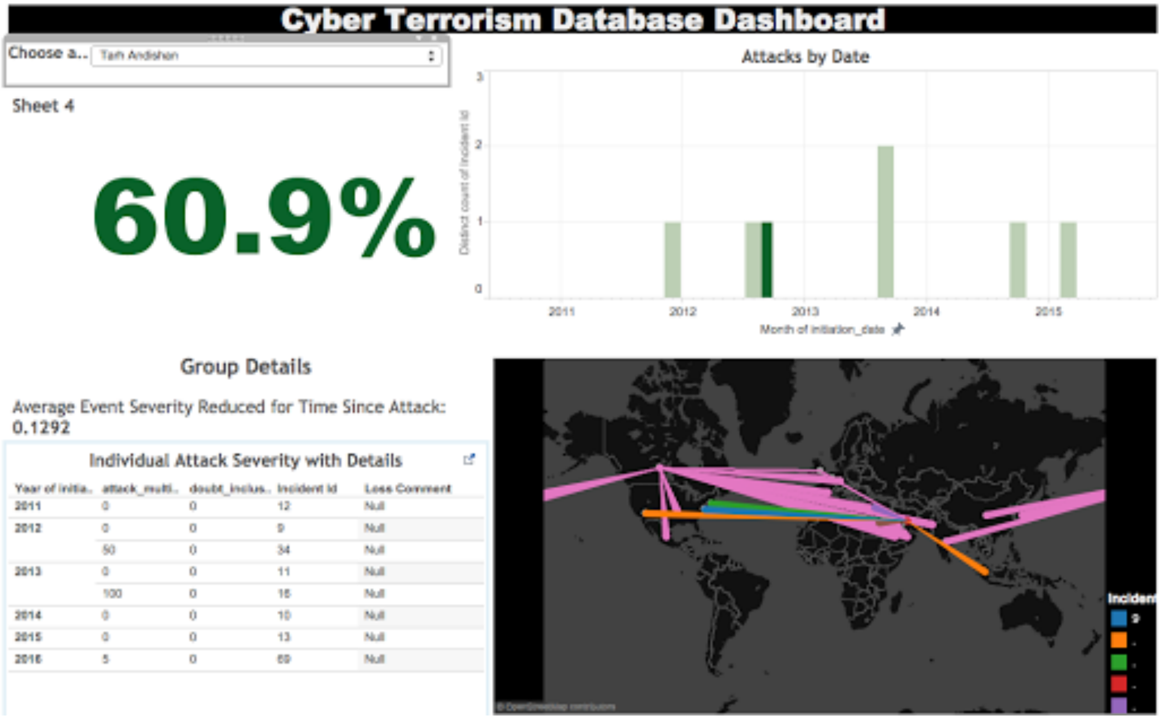
Group Details

Average Event Severity Reduced for Time Since Attack: 0.1292

Individual Attack Severity with Details

Year of initia..	attack_multi..	doubt_inclus..	Incident Id	Loss Comment
2011	0	0	71	Null
			72	Null
			73	Null
2012	0	0	74	Null
2013	0	0	76	Null
			80	Null
	2	0	77	Null
			79	Null
	7	0	78	Null
	16	0	75	Null

Figure 6: Syrian Electronic Army Dashboard Report



Cyber Terrorism Database Dashboard

Choose a.. Tarh Andishan

Sheet 4

60.9%

Group Details

Average Event Severity Reduced for Time Since Attack: 0.1292

Individual Attack Severity with Details

Year of initia..	attack_multi..	doubt_inclus..	Incident Id	Loss Comment
2011	0	0	12	Null
2012	0	0	9	Null
	50	0	34	Null
2013	0	0	11	Null
	100	0	16	Null
2014	0	0	10	Null
2015	0	0	13	Null
2016	5	0	69	Null

Figure 7: Tarh Andishan Dashboard Report

If these reports are compared, one finds that the SEA has both a greater number of attacks at a higher, more consistent frequency. This appears to be the deciding factor in assigning the SEA to be a greater threat than Tarh Andishan; however, this is too simplistic. What the database currently lacks is the ability to successfully differentiate between the severities of attacks that have *similar* propagation methods. For instance, most of the SEA attacks were access attacks into Twitter accounts. Tarh Andishan similarly used access as a propagation method when it hacked into the Rye Brook damn just north of NYC in 2013. Since these attacks utilized the same propagation methods, they are granted similar weights. However, this fails to recognize, and report, the reality of the situation. The latter attack breached a much higher level of security and had a far greater potential for physical disaster than did the hacking of a social media account.

Further complicating matters is that of the incident multiplier, which can be seen in the above reports. In the case of the SEA, all the incidents recorded in the database were reports on individual attacks that may or may not have had more than one victim. If an attack had multiple target victims, this was reported in the attack multiplier section. In the case of Tarh Andishan, the majority of the information gathered came from industry white papers that did not specify individual attacks, but rather mentioned the total number of attacks that were perpetrated over a period of time. For our purposes, and as a result of lack of time and resources, we coded these as a single incident with a high attack multiplier. This dramatically decreased the number of individual attacks attributed to the group.

This perceived discrepancy in the database's findings versus actuality reveals the need to reevaluate the algorithms used in order to produce results more indicative of reality. This can be fixed by hiring full-time data analysts and cyber security experts.

Conclusion: Creating A Sustainable Global Cyber Attacks Database

We believe the MECAdb in its present form proves the concept of open source data aggregation for understanding cyber actors in the Middle East. While it requires further development, we believe that given time and resources, this is a viable tool. The MECAdb's ultimate goal is to become the first analytic tool of its kind focused on the region. However, we hope that through our extended efforts, it will lead to other regional datasets and eventually be implemented on a global scale, becoming a Global Cyber Attacks Database (GCAD). However, in order for the current MECAdb to turn into an effective GCAD, the appropriate resources and manpower must be allocated. The GCAD would require a full-time team consisting of administrative staff, including directors, data scientists, software engineers, data collection managers, and cyber threat researches. Additionally, the GCAD must have access to advanced analytic tools to automate the data collection process in a similar way to the GTD.

While cyber security exists in a technical domain, its ramifications are multidisciplinary and extend beyond the cyber community. The GCAD has the potential to serve as a resource for both academic research and national-level security, in addition to being a viable educational tool for private firms looking to protect themselves against attacks.

In order to sustain the GCAD, it will require both capital and political investments. Currently, the private cyber security industry is valued at $77 billion and is expected to surpass the $100 billion mark by 2020.[20] In the private business realm, international corporations and private security firms have a vested interest in secure networks, both for

[20] Cyber Security Market Report Q4 2015. CyberSecurity Venture

themselves and their clients. Such groups have already invested funds toward better understanding and responding to cyber threats and are likely to continue to do so as cyber security becomes an ever-growing issue.

Politically, cyber security is quickly becoming a top priority. In 2013, US Computer Emergency Readiness Team (US-CERT) responded to 228,700 cyber incidents involving federal agencies, companies that run critical infrastructure, and contract partners. That number is nearly 120,000 greater than what it was in 2009. [21] The US government has recognized and responded to these attack by increasing efforts to thwart cyber attacks. The FY2016 Department of Homeland Security (DHS) states:

Cyber security is of growing relevance to our national and economic security. Funding in this request supports the Department's two flagship cyber acquisition programs—the National Cybersecurity Protection System and Continuous Diagnostics and Mitigation—which enhance cyber security situational awareness and information sharing

We believe that cyber security infrastructure must extend past the hard drives, servers and terminals just as international security must extend past the guns and bombs. We believe that understanding the persons and groups sitting in front of terminal are critical variables in combating the growing cyber threats in the Middle East and beyond and we believe we have taken the first step to understand these groups with our Middle East Cyber Attack Database.

[21] Marilym Cohodas. "Why We Need Better Cyber Security: A Graphical Snapshot." *Information Week.* November 28, 2014.

Bibliography

Al-Rawi, Ahmed. "Cyber warriors in the Middle East: The case of the Syrian Electronic Army ." *Public Relations Review* (Department of Media & Communication, School of History, Culture, & Communication, Erasmus University, Rotterdam, The Netherlands) 40 (April 2014): 420-428.

Al-Saggaf, Yeslam. "The Online Public Sphere in the Arab World: The War in Iraq on the Al Arabiya Website." *Journal of Computer-Mediated Communication* (International Communication Association) 12 (2006): 311-334.

Arreguin-Toft, Ivan. "How the Weak Win Wars: A Theory of Assymetric Conflict." *International Security*, June 1, 2011: 93-128.

Asal, Victor H., and R. Carl Rethemeyer. *Big Allied and Dangerous Dataset Version 2.* 2015. www.start.umd.edu/baad/database (accessed April 20, 2016).

Beauchesne, Ann. "Urgent need for cybersecurity legislation." *The Hill.* October 20, 2015. http://thehill.com/blogs/congress-blog/technology/257336-urgent-need-for-cybersecurity-legislation (accessed April 10, 2016).

Bureau of Counterterrorism. *Country Reports on Terrorism 2014.* Report, Bureau of Counterterrorism, United States Department of State, Washington, D.C.: United States Department of State, 2015.

Carr, Jeffrey. *Inside Cyber Warfare.* Sebastopol, California: O'Reilly Media, Inc, 2010.

Clarke, Richard A., and Robert K. Knack. *Cyber War: The NExt Threat to National Security and What to Do About It.* New York, New York: HarperCollins Publishers, 2010.

Cohodas, Marilym. "Why We Need Better Cyber Security: A Graphical Snapshot." *Information Week.* November 28, 2014. http://www.darkreading.com/operations/why-we-need-better-cyber-security-a-graphical-snapshot-/d/d-id/1317398 (accessed April 10, 2016).

Coleman, Gabriella. *Hacker, Hoaxer, Whistleblower, Spy: The Many Faces of Anonymous.* Verso, 2014.

CrowdStrike. *Global Threat Report 2015.* Crowdstrike, CrowdStrike, 2015.

Cybersecurity Ventures. "Cybersecurity Market Report Q3 2015." *Cybersecurity Ventures.* September 1, 2015. http://cybersecurityventures.com/cybersecurity-market-report/ (accessed December 2, 2015).

Debeck, Charles. "The Correlates of Cyber Warfare: A database for the modern era." *Graduate Theses and Dissertations* (Digital Repository @ Iowa State University) Paper 12062 (2011).

Falliere, Nicolas, Liam O. Murchu, and Eric Chien. *W32.Stuxnet Dossier.* Dossier, Symantec Security Response, Cupertino: Symantec Corporation, 2011.

Foster, Peter. "'Bogus' AP tweet about explosion at the White House wipes billions off US markets." *The Telegraph.* April 23, 2013. http://www.telegraph.co.uk/finance/markets/10013768/Bogus-AP-tweet-about-explosion-at-the-White-House-wipes-billions-off-US-markets.html (accessed April 2, 2016).

Gleditsch, Kristian S., and Michael D. Ward. "Interstate System Membership: A Revised List of the Independent States Since 1816." *International Interactions* 25 (1999): 393-413.

Goodchild, M.F. "Geographic information Systems: today and tomorrow." *Annals of GIS* 15, no. 1 (2009).

Goodchild, Michael F. "Rediscovering the World Through GIS." *National Center for Geographic Information and Analysis*, 1998.

Herr, Trey. *PrEP: A Framework for Malware and Cyber Weapons.* Thesis, Cyber Security and Policy Research Institute , George Washington University, Washington, DC: Cyber Security and Policy Research Institute, 2014.

Heydemann, Steven. *Upgrading Authoritarianism in the Arab World.* Analysis Paper, The Saban Center for Middle East Policy, The Brookings Institution, The Brookings Institution, 2007.

Hirschfeld Davis, Julie. "Hacking of Government Computers Exposed 21.5 Million People." *The New York Times.* July 9, 2015. http://www.nytimes.com/2015/07/10/us/office-of-personnel-management-hackers-got-data-of-millions.html (accessed December 2, 2015).

Institute for Economics & Peace. *Global Terrorism Index 2015: Measuring and Understanding the impact of Terrorism.* Index, Institute for Economics & Peace, http://economicsandpeace.org/wp-content/uploads/2015/11/Global-Terrorism-Index-2015.pdf, 2015.

Kaspersky Labs. *The Desert Falcons Targeted Attacks.* Report, Kaspersky Labs, Moscow: Kaspersky Labs, 2015.

Lafree, Gary, and Laura Dugan. "Introducing the Global Terrorism Database." *Terrorism and Political Violence* 19, no. 2 (July 2007): 181-204.

LaFree, Gary, Laura Dugan, Heather V. Fogg, and Jeffrey Scott. *Building a Global Terrorism Database.* Grant Report, University of Maryland, U.S. Department of Justice, unpublished, 2006.

Lowenthal, Mark M. *Intelligence: From Secrets to Policy.* Los Angeles, CA: SAGE/CQ Press, 2012.

Lynch, Marc. "Media, Old and New." In *The Arab Uprisings Explained: New Contentious Politics in the Middle East*, edited by Marc Lynch, 93-109. New York, New York: Columbia University Press, 2014.

Mack, Andrew J.R. "Why Big Nations Lose Small Wars: The politics of Asymmetric Conflict." *World Politics* 27, no. 2 (1975): 175-200.

National Consortium for the Study of Terrorism and Responses to Terrorism (START). "Global Terrorism Database." *Retrieved from http://www.start.umd.edu/gtd.* 2013.

National Consortium for the Study of Terrorism and Responses to Terrorism. *Global Terrorism Database.* June 1, 2015. http://www.start.umd.edu/gtd/ (accessed November 2015, 2015).

O'Day, Alan, ed. *Cyberterrorism.* Burlington, VT: Ashgate Publishing Company, 2004.

Office of the Director of National Intelligence. *Intelligence Community Home.* December 2, 2015. http://www.dni.gov/index.php (accessed December 2, 2015).

Olsen, Parmy. *We Are Anonymous: Inside the Hacker World of LulzSec, Anonymous and the Global Cyber Insurgency.* New York, New York: Litte, Brown and Company, 2013.

Reuters. "Yemeni group hacks 3,000 Saudi govt computers to reveal top secret docs." *RT.* May 22, 2015. https://www.rt.com/news/261073-yemen-cyber-hack-saudi/ (accessed April 10, 2016).

Rid, Thomas. *Cyber War Will Not Take Place.* New York, New York: Oxford University Press, 2013.

START. *Global Terrorism Database Codebook: Inclusion Criteria and Variables.* Codebook, Global Terrorism Database, National Consortium for the Study of Terrorism and Responses to Terrorism, College Park: National Consortium for the Study of Terrorism and Responses to Terrorism, 2015.

Stephens, Cory. "Electronic Armies: Hacktivism and Cyber Threat Actors in the Middle East." Unpublished, American University of Beirut, May 2015.

—. "Middle East Cyber Attack Database: Theoretical Foundations and Methodological Considerations." *Unpublished.* Washington, D.C.: Unpublished, December 7, 2015.

—. "Toward a Center for Middle East Cyber Analysis and Cyber Attack Databases." *Unpublished.* Washington, District of Columbia: Unpublished, George Washington University, Dec 2015.

The Guardian. "Turkey power outage shuts down public transportation and half of provinces." *The Guardian.* March 31, 2015. http://www.theguardian.com/world/2015/mar/31/turkey-power-outage-shuts-down-transportation-provinces (accessed April 10, 2016).

United States of America v. Ahmad Fathi, Hamid Firoozi, Amin Shokohi, Sadegh Ahmadzadegan, Omid Ghaffarinia, Sina Keissar, and Nader Saedi. (United States District Court Southern District of New York, March 24, 2016).

United States of America v. Ahmad 'Umar Agha and Firas Dardar. 1:14-MJ-292 (United States District Court for the Eastern District of Virginia, June 12, 2016).

United States of America v. Peter Romar and Firas Dardar. 1:15-mj-00498-MSn (United States District Court for the Eastern District of Virginia, September 29, 2015).

Zetter, Kim. "Sony Got Hacked Hard: What We Know and Don't Know So Far." *Wired.* December 3, 2014. http://www.wired.com/2014/12/sony-hack-what-we-know/ (accessed December 2, 2015).