Codebook for the Dyadic Cyber Incident and Dispute Data, Version 1.5: DRAFT

Ryan C. Maness
And
Brandon Valeriano

Suggested citations:

For Codebook:
Updated Data Collection Effort
Maness, Ryan C., Brandon Valeriano, and Benjamin Jensen. 2016. Coding Manual for v1.5 of the Dyadic Cyber Incident and Dispute Dataset, 2000-2014, Unpublished manuscript.

Published Version
Valeriano, Brandon and Ryan C. Maness. 2015. *Cyber War versus Cyber Realities: Cyber Conflict in the International System* (New York: Oxford University Press).
Book version has one additional dispute

(and)
Valeriano, Brandon and Ryan C. Maness. 2014. "The Dynamics of Cyber Conflict between Rival Antagonists, 2001-2011." *Journal of Peace Research,* 51 (3): 347-360.

Overview

This codebook presents a point of reference for variables for dyadic rival states that are in Cyber Conflict Data Project v 1.5 for the years 2000-2015 (Valeriano and Maness 2014, 2015). Rival dyads are extracted from the Klein, Diehl and Goertz (2006) enduring rival dataset as well as Thompson's (2001) strategic rival dataset. Each pair of states engaged in cyber conflict has two states involved, on opposite sides of the cyber incidents and disputes. For individual cyber conflicts, we use the phrase 'cyber incident.' Incidents may include thousands of intrusions, but accounting for every single intrusion the operation made is impossible and unwieldy. For operations containing a number of incidents that are part of an overall cyber campaign we use the term 'cyber disputes.'

Each of the two states directly involved in cyber incidents and disputes against the other (non-state actors or entities can be targets but not initiators as long as they critical to state based systems, or if the original hack escalates into an international incident in the non-cyber domain,) - the initiation must come from the government or there must be evidence that an incident or dispute was government sanctioned (see below for responsibility confirmation). For the target state, the target must be a government entity, either military or non-military; or a private entity that is part of the target state's national security apparatus (Power grids, defense contractors, and security companies), an important media organization, or multinational corporation. The dataset does not include multilateral cyber incidents or disputes; these types of incidents or disputes are only coded at the dyadic level. Third parties are noted and coded separately. Furthermore, a dispute ends after there has been a substantial period of time before the next, if the tactics by the initiator substantially changes, or it is clear the action has terminated.

Version 1.0 of DCID uncovered 126 active rival dyads in the data (Valeriano and Maness 2014, 2015). Valeriano and Maness (2014) found 110 cyber incidents within 45 overall disputes among 20 of the 126 pairs of states. (The 2015modificationin the OUP book contained 111 incidents among 45 disputes, where an espionage dispute between China and India was added).Version 1.5 has expanded to 165 incidents within 51 disputes from the years 2000-2014. It includes new variables and coding methods (detailed below), as well as expanding the inclusiveness of relevant non-state targets to include national security

contractors, media organizations, and other relevant multinational corporations such as banks, technology companies (Google, Apple), and utility companies for the years 2000 to 2014. We do not code non-state initiators in this dataset; the initiators must be state entities. Groups such as the Syrian Electronic Army, cyber-jihadists such as the Islamic State, or hacktivist groups such as Anonymous are not included as this would expand the purpose and scope of this data beyond measure.

This is an active dataset and will be maintained at cyberconflictdata.com as more cyber incidents and disputes are expected through time if support is available.

Specific Procedures

The Cyber Conflict Data Project was developed and written by Ryan C. Maness and Brandon Valeriano to develop a replicable and reliable dataset for all cyber incidents and disputes between states and relevant non-state targets.

The coding method specifically follows the Correlates of War procedures in examining sources throughout history, in the media, and, new for cyber conflict, from government or critical cyber security firm reports. An example of a Correlates of War dataset is the Militarized Interstate Disputes collection, which records cases of conflict between states "in which the threat, display or use of military force short of war by one member state is explicitly directed towards the government, official representatives, official forces, property, or territory of another state" (Jones, Bremer, and Singer 1996). It uses historical and diplomatic sources to isolate and codify each isolated incident. Cyber conflict is a more recent phenomenon than militarized disputes, as we demark the beginning of widespread international cyber conflict to begin with the year 2000. Therefore, we are able to access information on cyber incidents and disputes exclusively from the internet, using the Google News search engine as our uniform data extraction tool, as well as the other sources mentioned. In the future automatic events data searches may be undertaken but for now we are confident we can maintain an active dataset using focused search.

For the purposes of this study, electromagnetic pulses (EMPs), radar jamming, laser jamming/deception, and other measures/countermeasures traditionally considered electronic warfare (EW) are not defined as cyber incidents. Cyber incidents require the manipulation of computer code for malicious purposes. Electronic manipulation either damage or destroy circuitry through electronic (i.e. radio waves) and/or directed energy. An example of this would be the hacking of Unmanned Aerial Vehicles (UAVs) aka drones to bring them down. However, security for these vehicles has been much better, and this is a very rare phenomenon. While the two can overlap, we focus on cyber conflict as the manipulation of code through networks.

Given that searches can be conducted digitally, we focus on the following search terms to start our investigation. These search parameters are not exclusive and the coder should endeavor to examine computer security reports and government information where possible.

Enter the "Google News" search engine and enter "participant A eg. Iran" AND "participant B eg. Israel" AND "cyber" OR "internet attack" OR "infrastructure attack" OR "government cyber attack" OR "network breach" OR "hack" and customize the date range for 1/1/2000 to 12/31/2014.

After an incident is identified, computer security firm reports and government reports are used to further code each incident. Cyber security archive methods such as email digests, web based efforts, and Twitter accounts will be used to cull potentially missed technical reports.

What to look for and record:

A.. The dyad (states involved), only two states recorded per incident, see rules for third party coding in Hand I below
B. Start and end date of interaction, cyber-incident and dispute, most incidents will last a matter of days, weeks, or months and will be reported by the source used. If a specific end date cannot be found, please leave a note for the data managers. Disputes begin and end dates will be managed by the data managers.
C. Method of interaction/incident, 1-4 with decimal denotations for infiltrations (methods are listed below) for both incidents and disputes (see Sections I and VIII): defacements are vandalism; DDoS, zombies, botnets, and the like will be denial of service; any incident that uses spear phishing will be an intrusion, which includes Trojans, trapdoors, and backdoors. Intrusions are used in most theft/espionage operations; infiltrations, are usually worms or viruses, but can also be packet sniffers, logic bombs, and keystroke loggings; APTs can be any of the 4 methods; but must hit specific targets and be relatively undetectable for a time. If you are unsure as to whether or not an incident is an APT, please leave a not for the data manager.
D. Type of interaction (nuisance, defensive, offensive) for both incidents and disputes (see Sections II and VI)
E. The type of target (private/non-state, government non-military, government military, See Sections III and VII)
F. The initiator of the interaction (use COW country codes, when it is a two-way operation, enter both country codes, See Section XII)
G. The foreign policy motive of the initiator for disputes only (See Section IX)
H. The specific political objective of the cyber incident (See Section X).
I. Whether or not the political objective changed the behavior of the target state.
J. Whether or not a third party was involved in the initiation (other state, rebel group, corporation) 1 = yes, 0 = no; Sometimes, but not often, third party states will be involved in the initiation of a cyber incident. Look for explicit evidence that a third party was involved. Israel was a part of the United States' Stuxnet operation, for example.
K. Whether or not a third party was a target of the interaction 1= yes, 0 = no: This is more commonplace, especially for theft and espionage campaigns (intrusions)
L. Whether or not an official government statement was issued by the initiator, 0= no comment, 1= denial, 2= acceptance, 3-multiple; this will help in the responsibility coding; although most of the time governments will deny or not comment about their part in cyber incident initiation.
M. Damage level on the 0-10 scale level, given below for both incidents and disputes (same procedure for both, See Section IV)
N. Damage type (1. Direct and immediate, 2.Direct and delayed, 3.Indirect and immediate, 4.Indirect and delayed) See Section V.
O. The main source for the cyber interaction, paste all links to sources here
P. Any special notes pertaining to the interaction, write a short summary as to the particulars of the cyber incident

Once these procedures are finished, responsibility is the next and very important step in the coding process. To verify that the initiator was in fact the government or a government-sanctioned activity, the coding process goes through another process of verification. Attribution of cyber incidents and disputes can be a problematic issue; therefore we focus on what is called responsibility (Goodman 2010:128).One of the advantages of a cyber dispute is deniability. In our dataset, states that use information warfare must be fairly explicit and evident. If the responsibility of a dispute is in serious doubt, we do not code it as a state-based action. We do not take conventional wisdom at its word for operations and instead analyze the history of relations, the intent of the tactic, likelihood of government complacency and code disputes from this perspective. Therefore, simple news stories extracted by search engines such as "Google News" are not enough to make the dataset. Responsibility must be verified by government statements, policy reports, internet security firm reports, white papers from software security firms (Symantec, McAfee, Kapersky), or cyber-security magazines.

Coding for isolated incidents: For individual cyber conflicts, we use the phrase 'cyber incident.' Incidents such as Shady Rat include thousands of intrusions, but accounting for every single intrusion the operation made is impossible and unwieldy. Therefore, Shady Rat and other multiple-intrusive incidents are coded as just one incident per dyad as long as the goals and perpetrators remain stable. Each cyber incident is directed by one state or on behalf of the state against another state or state's national security apparatus or relevant multinational corporations.

I. Methods of cyber-incidents

Many news sources will report cyber incidents as viruses, because they do not have the technical know-how to categorize these types of interactions. It is important that coders are aware of this and make sure to code these incidents properly by finding reports on cyber incidents from government statements, policy reports, internet security firm reports, white papers from software security firms (Symantec, McAfee, Kapersky), or cyber-security magazines. The news search is the primer to find cyber incidents; the latter documents are what you will need to code these incidents properly.

1. Vandalism: Website defacements: Hackers use SQL injection or cross-site scripting (forms of command code) to deface or destroy victims' web pages. Although rather benign, these attacks may have important psychological effects.
2. Denial of Service: DDoS, distributed denial of service: DDoS attacks flood particular Internet sites, servers, or routers with more requests for data than the site can respond to or process. The effect of such an attack effectively shuts down the site thus preventing access or usage. Government sites important to the functioning of governance are therefore disrupted until the flooding is stopped or the attackers disperse. Such attacks are coordinated through "botnets," or a network of computers that have been forced to operate on the commands of an unauthorized remote user. The primary impact of DDoS attacks via botnets is the temporary disruption of service.
3. Intrusion: "Trapdoors" or "Trojans" and Backdoors: Trapdoors or Trojans are unauthorized software added to a program to allow entry into a victim's network or software program to permit future access to a site once it has been initially attacked. The purpose of trapdoors is to steal sensitive information from secured sites. Spear phishing is utilized to inject these cyber methods into networks. Here the initiator sends emails to employees or contractors of the targeted network, and if the email is opened, the intrusion is introduced to the system. The botnet technique is another option where a human being injects the intrusion from a portable drive such as a USB or disk.
4. Infiltration: Examples of attacks include logic bombs, viruses, packet sniffers, and keystroke logging. These methods force computers or networks to undertake tasks that they would normally not undertake. 1)Logic bombs are programs that cause a system or network to shut down and/or erase all data within that system or network. 2)Viruses are programs which attach themselves to existing programs in a network and replicate themselves with the intention of corrupting or modifying files. 3)Worms are essentially the same as viruses, except they do not need to attach themselves to existing programs. 4)Keystroke logging is the process of tracking the keys being used on a computer so that the input can be replicated in order for a hacker to infiltrate secure parts of a network.

   General infiltrations, packet sniffers or beacons, are not coded in this dataset, as most of the time no act of cyber malice is committed. They are monitoring techniques that search for certain information coming through internet cellular channels. If a potential incident is labeled as a packet sniffer or beacon, do not code it.

When infiltration is found, please try to delineate the type and decimal the number with the 4 (.1 logic bombs, .2 virus, .3 worm, .4 keystroke logging)

Note about APTs for incidents: APTs are different from traditional targeted methods in that they are customized, move slower to avoid detection, their intentions usually are more malicious and advanced and almost certainly come from states, and their targets are much more specific. Because all methods can also be Advanced Persistent Threats, APTs, we will code them with a dummy of "1" after the designated type

II. Interaction type

1. Nuisance (probing, disruption, chaos); most vandalism and denial of service incidents, intent is disrupt the day to day operations of a network, easily removable by target
2. Defensive operation (Cisco Raider, Buckshot Yankee, Israeli operations against cyber-jihad); the initiator must be the victim of a cyber incident first; these are defensive measures launched by a target where it becomes the initiator
3. Offensive strike (Ghost Net, Shady RAT, Stuxnet); intent is usually theft or espionage or to disrupt a specific national security strategy of a target, most intrusions and infiltrations.

III. Target type

1. Private/non-state (financial sector, power grid, defense contractor, media organization, MNC)
2. Government non-military (US State Department, government websites, government member website)
3. Government military (US Defense Department, US Cyber Command, US Strategic Command)

IV. Severity scale: the same for incidents and disputes, for disputes, the highest incident is the severity code for the dispute

10-Massive death as a direct result of cyber incident

Example - NORAD hacked and missiles launched, Air traffic control systems manipulated, commercial airliner hacked and brought down

Notes - For this measure to be coded, a state must direct a cyber incident against another state's or private organizations' network where the system is manipulated and massive loss of life is a result (over 100 deaths).

9- Critical national infrastructure destruction as a result of cyber incident

Example - power grid hack, hydroelectric dams shut down, indirect death

Notes - For this measure to be coded, a state's critical infrastructure must be breached and the network manipulated so that widespread functionality is disrupted for a period of time.

8-Critical national economic disruption as a result of cyber incident

Example - stock market price manipulation, critical e-commerce shut down for extended periods

Notes - For this measure to be coded, a sophisticated infiltration must be responsible for the manipulation of prices that affect stock market indexes and prices for extended periods of time. Another example would be a cyber incident being responsible for the slowing or shutting down commerce online. This attack must be severe and critically threatening beyond compromising payment systems.

7-Minimal death as direct result of cyber incident

Example - Auto hacked, pacemaker hacked

Notes - Here a state-sponsored cyber incident would be responsible for the death of an individual or group of individuals of another state by either hacking into the automobile of the victim(s) or causing it to crash, or if the victims(s) are dependent on a pacemaker to live and this device is hacked, leading to that person's death.

6-Single critical network widespread destruction

Example - (Shamoon, DoD taken offline, Lockheed Martin database wiped out)

Notes - For this measure to be coded, a single network that is critical to national security must be breached and widespread destruction must be successful. Critical stored information is destroyed or unrecoverable or functionality of the network must be limited to non-existent for a period of time.

5-Single critical network and physical attempted destruction

Example - (Stuxnet, Flame, DoD secure network intrusion)

Notes - This measure entails the successful breach of a network where damage is done, however the breached network is left intact in terms of functionality and recoverable losses.

4-Widespread government, economic, military, or critical private sector theft of information

Example - (US OPM hack, DoD employee records stolen, IRS hack)

Notes –Phishing and intrusion espionage campaigns that successfully steal large troves of critical information, such as the OPM hack.

3-Stealing targeted critical information

Example - (Chinese targeted espionage, government-sanctioned cyber crime, Sony Hack)

Notes - This involves the use of intruding upon a secure network and stealing sensitive or secret information. The theft of Lockheed Martin's F-35 jet plans or the U.S. Department of Defense's strategy in the Far East are examples. Or if the target was critical to national security or the objective of the attack had national security implications. The piggy-back method is another example of this severity type. The United States' NSA was able to piggy back on China's Byzantine Series undetected and spy on the targets that the original espionage was spying upon.

2-Harrassment, propaganda, nuisance disruption

Example - (Propagandist messages in Ukraine, Vandalism, DDoS in Georgia, Bronze Soldier dispute)

Notes–Mainly vandalism or DDoS campaigns, this measure is coded when pockets of government or private networks are disrupted for periods of time and normal day to day online life is difficult, but recoverable.

1-Probing without kinetic cyber

Example - (US NSA dormant infiltrations)

Notes - Using cyber methods to breach networks but not utilize any malicious actions beyond that. Hacking a power grid but not shutting it down, planting surveillance technology within networks, and unsophisticated probing methods are examples of this severity level.

0-No cyber activity

V.Damage type(conceptualized from Rid and Buchanan 2014)

1. Direct and immediate: The term direct in this context means that the damage done by the cyber incident was what was intended by the initiator and the costs of the cyber incident are felt immediately. The Russian DDoS attacks on Estonia's government and private networks in 2007 is an example, as the effective shutdowns cost millions of dollars in lost revenue for the Baltic country.
2. Direct and delayed. Stuxnet was intended to disrupt Iran's nuclear program by damaging the centrifuges at the Natanz plant, and it succeeded. The impact of this attack took a number of months if not years to slowly disrupt and damage these centrifuges through code manipulation.
3. Indirect and immediate. Indirect in this context means that the damage done by the cyber incident was not the original intent of the initiator. The stealing of confidential information from a bank or a breach in the Wall Street system is an example of this. The costs of these incidents are felt immediately. Reputational damage or loss of confidentiality is what to look for when coding this damage.
4. Indirect and delayed. If intellectual property is stolen by an initiator and it becomes publicly available, this may result in improved competition for states or private companies that did not have this technology or advantage prior. China stole the American company's F-35 jet plans, and if it gave these plans to Russia, the effects of this cyber incident would be indirect and the costs would be felt at a future point in time.

Coding for cyber disputes (For Data Managers only, coders only code cyber incidents) For operations containing a number of incidents that are part of an overall cyber campaign we use the term 'cyber disputes.' For example, incidents such as GhostNet, Shady Rat, the Pentagon Raid, and the F-35 jet plan theft initiated by China against the United States and the American responses of Buckshot Yankee and Cisco Raider are all part of one sustained cyber dispute between the two rivals. Cyber disputes may contain only one incident or dozens. Furthermore, the initiator of the dispute or incident must be from a government or government affiliates in order for an operation to be included in our dataset. Targets may be non-state if they are important to a state's national security. Lockheed Martin, Mitsubishi, large banks, and Boeing are examples of non-state targets relevant to the national security of a state.

VI. Interaction type

1. Nuisance
2. Defensive operation
3. Offensive strike
4. Nuisance and defensive
5. Nuisance and offensive
6. Defensive and offensive
7. Nuisance, defensive, and offensive all involved

VII. Target type

1. Private/non-state but important to national security (financial sector, power grid, defense contractor)
2. Government non-military (state dept, govt websites, govt member website)
3. Government military (defense dept, cyber command, strategic command)
4. Private and government non-military
5. Private and government military
6. Government non-military and government military
7. Private, government non-military and government military

VIII. Methods for disputes

1. Vandalism
2. Denial of service
3. Intrusion
4. Infiltration
5. Vandalism and Denial of service
6. Intrusion and Infiltration

IX. Objectives for initiators

1. Disruption, (take down websites, disrupt online activities, Example: China's takedown of the New York Times and Washington Post in 2012 and Iran's Shamoon against Saudi Arabia's Aramco)
2. Theft/Espionage (steal sensitive information or strategies, Example: China's theft of Lockheed Martin's F-35 plans)
3. Coercion (abandon nuclear program, withdraw troops, Example: USA's Stuxnet against Iran; create chaos in a country to invoke a foreign policy response)

X. Specific political objective

Here we decipher as to why the cyber incident was launched in the first place. For example, for the Sony Hack the objective was to stop the release of the movie *The Interview*. A maximum of two political objectives are allowed.

XI. Did the political objective work?

Did the objective of the initiator work, i.e., did the target state change its behavior as a result of the cyber incident as a result? The Sony Hack did not work, as *The Interview* was released.

XII. Variables for the Dyadic Cyber Incident and Dispute (DCID) Dataset, Version 2

Variable Number	Variable Name	Variable Description
1	Cyberdisputenum	Cyber dispute number (Data managers only)
2	Cyberincidentnum	Cyber incident number (Data managers only)
3	DyadPair	State Pair ID (COW codes)
4	StateA	First state in dyad
5	StateB	Second state in dyad
6	Name	Name of cyber incident or dispute: most of the time the press or a security company gives the incident a name. If there isn't one, name the incident based on the anatomy of the attack
7	interactionstartdate	Cyber incident or dispute start date
8	interactionenddate	Cyber incident or dispute end date
9	interactiontype	Type of cyber interaction for incidents and disputes 1- Nuisance 2- Defensive operation 3- Offensive strike 4- Nuisance and defensive (disputes only) 5- Nuisance and offensive (disputes only) 6- Defensive and offensive (disputes only) 7- Nuisance, defensive, and offensive (disputes only)
10	Method	Cyber method utilized 1- Vandalism 2- Denial of Service (DDoS) 3- Intrusion 4- Infiltration 4.1 - Logic bomb 4.2 - Virus 4.3 - Worm 4.4 – Keystroke logging 5- Vandalism and Denial of Service (disputes only) 6- Intrusion and Infiltration (disputes only)
11	APT	Advanced Persistent Threat? 1- Yes, 0- No
12	Targettype	Type of target by cyber incident or dispute 1- Private/non-state 2- Government non-military 3- Government military 4- Private and government non-military (disputes only) 5- Private and government military (disputes only) 6- Government non-military and government military (disputes only) 7- Private, government non-military and government military (disputes only)
13	initiator	State that initiated the incident or dispute (COW code)

14	initiator objective	Objective of the initiating state (disputes only) 1- Disruption 2- Theft/Espionage 3- Coercion
15	Political objective	Statement of political objective of initiator
16	objective success	Did the objective work? 1- Yes, 0- No
17	3rdpartyinitiator	Third party involved with initiating state? 1- Yes, 0- No
18	3rdparty target	Third party involved as a target? 1- Yes, 0- No
19	Govtstatement	Statement from the initiating state? 0- No comment, 1- Denial, 2- Acceptance, 3- Multiple statements
20	Severity	Severity level of incident or dispute (for disputes code the highest incident severity 1- Probing without kinetic cyber 2- Harassment, propaganda, nuisance disruption 3- Stealing targeted critical information 4- Widespread government, economic, military or critical private sector theft of information 5- Single critical network and physical attempted destruction 6- Single critical network widespread destruction 7- Minimal death as a direct result of cyber incident 8- Critical national economic disruption as a result of cyber incident 9- Critical national infrastructure destruction as a result of cyber incident 10- Massive death as a direct result of cyber incident
21	Damage type	1. Direct and immediate 2. Direct and delayed 3. Indirect and immediate 4. Indirect and delayed
22	Source	The news source for the cyber interaction
23	Notes	Any special notes pertaining to the interaction

XIII.COW Country Codes

StateAbb	CCode	StateNme
USA	2	United States of America
CAN	20	Canada
BHM	31	Bahamas
CUB	40	Cuba
CUB	40	Cuba
HAI	41	Haiti
HAI	41	Haiti
DOM	42	Dominican Republic
DOM	42	Dominican Republic
JAM	51	Jamaica
TRI	52	Trinidad and Tobago
BAR	53	Barbados
DMA	54	Dominica
GRN	55	Grenada
SLU	56	St. Lucia

SVG	57	St. Vincent and the Grenadines
AAB	58	Antigua & Barbuda
SKN	60	St. Kitts and Nevis
MEX	70	Mexico
BLZ	80	Belize
GUA	90	Guatemala
HON	91	Honduras
SAL	92	El Salvador
NIC	93	Nicaragua
COS	94	Costa Rica
PAN	95	Panama
COL	100	Colombia
VEN	101	Venezuela
GUY	110	Guyana
SUR	115	Suriname
ECU	130	Ecuador
PER	135	Peru
BRA	140	Brazil
BOL	145	Bolivia
PAR	150	Paraguay
PAR	150	Paraguay
CHL	155	Chile
ARG	160	Argentina
URU	165	Uruguay
UKG	200	United Kingdom
IRE	205	Ireland
NTH	210	Netherlands
NTH	210	Netherlands
BEL	211	Belgium
BEL	211	Belgium
LUX	212	Luxembourg
LUX	212	Luxembourg
FRN	220	France
FRN	220	France
MNC	221	Monaco
LIE	223	Liechtenstein
SWZ	225	Switzerland
SPN	230	Spain
AND	232	Andorra
POR	235	Portugal
HAN	240	Hanover
BAV	245	Bavaria
GMY	255	Germany
GMY	255	Germany
GFR	260	German Federal Republic
GDR	265	German Democratic Republic
BAD	267	Baden
SAX	269	Saxony
WRT	271	Wuerttemburg
HSE	273	Hesse Electoral

HSG	275	Hesse Grand Ducal
MEC	280	Mecklenburg Schwerin
POL	290	Poland
POL	290	Poland
AUH	300	Austria-Hungary
AUS	305	Austria
AUS	305	Austria
HUN	310	Hungary
CZE	315	Czechoslovakia
CZE	315	Czechoslovakia
CZR	316	Czech Republic
SLO	317	Slovakia
ITA	325	Italy
PAP	327	Papal States
SIC	329	Two Sicilies
SNM	331	San Marino
MOD	332	Modena
PMA	335	Parma
TUS	337	Tuscany
MLT	338	Malta
ALB	339	Albania
ALB	339	Albania
MNG	341	Montenegro
MAC	343	Macedonia
CRO	344	Croatia
YUG	345	Yugoslavia
YUG	345	Yugoslavia
BOS	346	Bosnia and Herzegovina
KOS	347	Kosovo
SLV	349	Slovenia
GRC	350	Greece
GRC	350	Greece
CYP	352	Cyprus
BUL	355	Bulgaria
MLD	359	Moldova
ROM	360	Romania
RUS	365	Russia
EST	366	Estonia
EST	366	Estonia
LAT	367	Latvia
LAT	367	Latvia
LIT	368	Lithuania
LIT	368	Lithuania
UKR	369	Ukraine
BLR	370	Belarus
ARM	371	Armenia
GRG	372	Georgia
AZE	373	Azerbaijan
FIN	375	Finland
SWD	380	Sweden

NOR	385	Norway
NOR	385	Norway
DEN	390	Denmark
DEN	390	Denmark
ICE	395	Iceland
CAP	402	Cape Verde
STP	403	Sao Tome and Principe
GNB	404	Guinea-Bissau
EQG	411	Equatorial Guinea
GAM	420	Gambia
MLI	432	Mali
SEN	433	Senegal
BEN	434	Benin
MAA	435	Mauritania
NIR	436	Niger
CDI	437	Ivory Coast
GUI	438	Guinea
BFO	439	Burkina Faso
LBR	450	Liberia
SIE	451	Sierra Leone
GHA	452	Ghana
TOG	461	Togo
CAO	471	Cameroon
NIG	475	Nigeria
GAB	481	Gabon
CEN	482	Central African Republic
CHA	483	Chad
CON	484	Congo
DRC	490	Democratic Republic of the Congo
UGA	500	Uganda
KEN	501	Kenya
TAZ	510	Tanzania
ZAN	511	Zanzibar
BUI	516	Burundi
RWA	517	Rwanda
SOM	520	Somalia
DJI	522	Djibouti
ETH	530	Ethiopia
ETH	530	Ethiopia
ERI	531	Eritrea
ANG	540	Angola
MZM	541	Mozambique
ZAM	551	Zambia
ZIM	552	Zimbabwe
MAW	553	Malawi
SAF	560	South Africa
NAM	565	Namibia
LES	570	Lesotho
BOT	571	Botswana

SWA	572	Swaziland
MAG	580	Madagascar
COM	581	Comoros
MAS	590	Mauritius
SEY	591	Seychelles
MOR	600	Morocco
MOR	600	Morocco
ALG	615	Algeria
TUN	616	Tunisia
TUN	616	Tunisia
LIB	620	Libya
SUD	625	Sudan
SSD	626	South Sudan
IRN	630	Iran
TUR	640	Turkey
IRQ	645	Iraq
EGY	651	Egypt
EGY	651	Egypt
SYR	652	Syria
SYR	652	Syria
LEB	660	Lebanon
JOR	663	Jordan
ISR	666	Israel
SAU	670	Saudi Arabia
YAR	678	Yemen Arab Republic
YEM	679	Yemen
YPR	680	Yemen People's Republic
KUW	690	Kuwait
BAH	692	Bahrain
QAT	694	Qatar
UAE	696	United Arab Emirates
OMA	698	Oman
AFG	700	Afghanistan
TKM	701	Turkmenistan
TAJ	702	Tajikistan
KYR	703	Kyrgyzstan
UZB	704	Uzbekistan
KZK	705	Kazakhstan
CHN	710	China
MON	712	Mongolia
TAW	713	Taiwan
KOR	730	Korea
PRK	731	North Korea
ROK	732	South Korea
JPN	740	Japan
JPN	740	Japan
IND	750	India
BHU	760	Bhutan
PAK	770	Pakistan
BNG	771	Bangladesh

MYA	775	Myanmar
SRI	780	Sri Lanka
MAD	781	Maldives
NEP	790	Nepal
THI	800	Thailand
CAM	811	Cambodia
LAO	812	Laos
DRV	816	Vietnam
RVN	817	Republic of Vietnam
MAL	820	Malaysia
SIN	830	Singapore
BRU	835	Brunei
PHI	840	Philippines
INS	850	Indonesia
ETM	860	East Timor
AUL	900	Australia
PNG	910	Papua New Guinea
NEW	920	New Zealand
VAN	935	Vanuatu
SOL	940	Solomon Islands
KIR	946	Kiribati
TUV	947	Tuvalu
FIJ	950	Fiji
TON	955	Tonga
NAU	970	Nauru
MSI	983	Marshall Islands
PAL	986	Palau
FSM	987	Federated States of Micronesia
WSM	990	Samoa

XIV. Coding Process Example 1: Shamoon

Link to documents where information for coding was extracted:
1. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6578789,
2. http://www.isssource.com/iran-behind-shamoon-attack/

Columns A.-B. : Leave blank for data managers.

C. Code the DyadPair column with the country code with the lesser value first, greater value second. For this case, the dyadic pair number is 630670.

Search for the countries involved; the initiator and dispute. For this case we find that Iran is the initiator and Saudi Arabia, via the state-owned oil company Aramco, is the target.

The target is not only state-owned, but oil is very important to Saudi Arabia so threats against its oil company threaten its national security. Paragraph 2 of the introductory section of the article reads, “The market value of this oil giant has been estimated at up to $10 trillion USD in some financial journals, making it the world's most valuable company [4]. Threats against Aramco could potentially jeopardize the national security of Saudi Arabia.” This is why we code it.

How do we know that the government of Iran is responsible for Shamoon?

What is the international context of Shamoon?

First, we know that the United States is a close ally of Saudi Arabia. Second, we know that Iran, still reeling from the damage of Stuxnet two years prior, wants retribution. Launching a cyber incident such as Shamoon against the US or a US oil company directly may invoke an escalated response from a much more powerful adversary, so launching the virus upon Saudi Arabia is the next best choice, yet is still a manageable incident for Iran. So how do we know it was an Iranian government sanctioned virus?

Looking at the first article, in the second paragraph of Section II, Part B, "Stuxnet has been described as 'the future of cyberwar' [12]. In response, Iran has developed an official Cyber-warfare Division under the Islamic Revolution Guards Corps. In 2011, Iran promised, "Should the Iranian cyber army be provoked, Iran would combat these operations with their own 'very strong' defensive capabilities" [13]. The country has also allegedly begun investing over $1 billion USD in improving its cyber-security defenses."

Then in the first paragraph of Section II, Part C, it notes, "Lessons for Saudi Arabia can be drawn from Iran's experience of cyber-warfare, and are especially pertinent since Iran is Saudi Arabia's primary regional rival [3]. However, public English and Arabic sources discussing Saudi cyber security are limited, with the exception of a few media mentions and an article authored by Brigadier General and member of the Saudi royal family, Prince Naef Bin Ahmed Al-Saud. Al-Saud did not provide many details about the state of cyber-security in the Kingdom. However, he implied that Saudi Arabia currently does not provide financial incentives to invest in cyber-security nor is there coordination between the Ministry of Defense and critical infrastructure stakeholders regarding cyber-defense."

What we know thus far is that Iran has been developing offensive cyber capabilities that it is looking to respond to Stuxnet, and that Saudi Arabia was vulnerable to these types of attacks in 2012. Furthermore, Iran and Saudi Arabia are better Middle Eastern regional rivals.

Now moving on to the second article in the links provided above, it reads (beginning with the third paragraph), "Two former senior CIA officials first alerted *ISSSource* the culprit in the attack was Iran working with personnel inside the Aramco's computer center. They said the Saudi regime is investigating the attack and is arresting suspects like operating staff, janitors, office people, and cargo handlers. CIA sources said attack was the work of a disgruntled Shiite insider (or insiders) that had full access to the system. Richard Stiennon at IT-Harvest, a firm that tracks and reports on evolving cyber threats, told *ISSSource* 30,000 computers ended up scrambled and Iran was the perpetrator. He said Iranian-trained hackers launched the attack "in deep wrath" because of the mistreatment of the Shiites at the facility, and in Syria and Bahrain — two countries where the Saudi government has reportedly aided Sunni factions in their struggle with the Alawite-dominated regime and the Shiite majority, respectively."

Responsibility for the Iranian government's complacency for Shamoon is now confirmed.

D.-E. Write the two countries' involved in the cyber incident, in order of COW country code. In this case, Iran is first, Saudi Arabia is second.

F. Write the name of the cyber incident. Most of the time the press or a security company gives the incident a name. If there isn't one, name the incident based on the anatomy of the attack, example: Iran Twitter Hack. In this case, Shamoon is the agreed-upon name.

G.-H. Interaction start and end dates. All sources on Shamoon report its begin date as August 15. 2012, so write 8/15/2012 in Column G. As the second report on this incident indicates, Shamoon did its work quickly, as when it was injected; it immediately went to work and erased all data in 30,000 computers within the Aramco network. Therefore, it ended almost as quickly as it started. We code the end date for the following day, 8/16/2012.

I. What type of interaction was Shamoon? According to the coding rules, Shamoon is an offensive strike because its intent was to knock out a country's specific national security strategy. Saudi oil, as the first report indicates is crucial to Saudi national security. Therefore, we write 3 in Column I.

J. The first report calls it a virus. They are not technical experts and are wrong. As we see in the second report, which is a computer security firm, they call it a Trojan Horse, and they are correct. Shamoon is an intrusion where the botnet method of entering the Aramco network was utilized. As the second article reports, the perpetrator injected the Trojan Horse with a USB drive, where this allowed remotely controlled computers to takeover all operations in the network, inject wiper malware, and erase all data. We write a 3 in Column J.

K. Was this an Advanced Persistent Threat? No. Although the attack had a specific target, the technology of Shamoon was relatively low and did not target anything specific in the Aramco network. It was easily detected and did its work quickly. APTs target specific parts of networks and move slowly to avoid detection. We therefore write a 0 in Column K.

L. What type of target is Aramco? We code this target as 2, government non-military because Aramco is majority owned by the Saudi government. If this happened to Exxon-Mobil, for example, we would code it as 1, a private company.
M. Who was the initiator? It is clear that Iran was the initiating state. We write its country code, 630, in Column M.

N. What was the objective of the initiator? Usually Trojans are used to steal information. Not in the Shamoon case. The intrusion was to wipe out all data in the Aramco network. Therefore, we do not code it as 2, theft/espionage, nor do we code it as 3, to coerce a state, because it is clear that Saudi Arabia would never stop producing oil as a result of Shamoon. It is therefore a disruption, and we code it as 1 in Column N.

O. What was the political objective of Shamoon? The objective was to stop all of Aramco's oil production as a result of the cyber incident. This is what we write in Column O.

P. Did Shamoon's political objective work? No. Although there was mass confusion and the network of Aramco was inoperable for a time, oil production, although slowed, continued. We write 0 in Column P.

Q. Was there a 3rd party that helped Iran initiate Shamoon? All reports on Shamoon report Iran as the only initiator. We write 0 in Column O.

R. Did Shamoon target another country? No, it is apparent that Saudi Arabia was the sole target. A few weeks later the Qatari natural gas company RasGas gets hit with a Shamoon-like attack, but it is separate from Shamoon. We write 0 in Column P.

S. How did Iran respond to Shamoon when it was accused of launching it? All documents report that Iran denied its part in the incident. We therefore write 1 in Column Q.

T. How do we rank Shamoon on the severity scale? Shamoon is clearly a 6 on the severity scale; it targeted a specific infrastructure of a country. It did not merely deface users screen nor did it just shut down the system for a few hours. It erased all digital data. It was therefore too severe to be coded as 5. Furthermore, it did not have a dramatic effect on the country of Saudi Arabia, nor were there any deaths. If all Saudi oil production was halted for a significant period of time, we would then be able to give it an 8 ranking. However, Saudi Arabia was able to produce oil during 8/15-16/2012. It lost sensitive data however, and Shamoon is therefore coded as 6 on the severity scale.

U. What type of damage did Shamoon inflict? As soon as the intrusion entered Aramco's network it began to erase data. Its effects were felt immediately. It also did what its initiators wanted it to do, which was erase data in Aramco's system. It succeeded. We therefore code Shamoon as a direct and immediate incident, and write 1 in Column U.

V. We copy and paste all sources we used to code Shamoon in this column. We used 2 for this case.

W. In this column we write a short summary of Shamoon. Here is the example: "Saudi Arabia's national oil company, Aramco, said on Sunday that a cyber attack against it in August that damaged some 30,000 computers was aimed at stopping oil and gas production in Saudi Arabia, the biggest exporter in the Organization of the Petroleum Exporting Countries."

Coding Process Example 2: Sony Hack

Link to documents where information for coding was extracted:
1. https://www.sans.org/reading-room/whitepapers/casestudies/case-study-critical-controls-sony-implemented-36022?utm_content=buffer69ebd&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer
2. https://www.riskbasedsecurity.com/2014/12/a-breakdown-and-analysis-of-the-december-2014-sony-hack/

Columns A-B: _____

C. Code the DyadPair column with the country code with the lesser value first, greater value second. For this case, the dyadic pair number is _____.

Search for the countries involved, the _____ and _____. For this case we find that _____ is the initiator and the United States, via the multinational corporation _____, is the target.

The reason we code the Sony hack in this case is because it leads to an international dispute in the _____domain. The international community nearly unanimously suspected the North Korean government as the culprit of the attack, as the row over the soon to be released, movie, *The Interview*, a satirical movie with leader Kin Jung-Un as the antagonist who is assassinated in the movie, was about to be released worldwide. The United States then responded to the attack by placing further economic sanctions on North Korea as a direct result of the cyber incident. This cyber incident, therefore, escalated into an international incident in the non-cyber domain.

How do we know that the government of North Korea is responsible for the Sony Hack?

What is the _____of the Sony Hack?

First, it seems that the only person and/or government that would take offense to *The Interview* are Kim Jung Un and North Korea. As Un is relatively new to the annals of power in North Korea, foreign policy experts are still trying to figure out how the new leader "ticks." Would he be willing to launch a major cyber attack over a slapstick comedy that he deems offensive? The group claiming responsibility called itself the _____ (P. 2 of first article). A spokesperson for the North Korean foreign ministry even went as far to declare that the movie's release would be considered an _____ (p. 3 of first article). IT can be safely assumed, therefore, that Kim Jung-Un was deeply offended in some way by the movie.

Looking at pp. 3-4 of the first article, on December 8, 2014, the United States informally blamed the North Korean government for the attack, and the FBI formally attributed North Korea on December 19. Therefore, as the attacker is an attributed government, we move forward with coding the Sony hack.

D.-E. Write the two countries' involved in the cyber incident, in order of COW country code. In this case, _____is first, _____is second.

F. Write the name of the cyber incident. Most of the time the press or a security company gives the incident a name. If there isn't one, name the incident based on the anatomy of the attack, example: Iran Twitter Hack. In this case, _____is the agreed-upon name.

G.-H. Interaction start and end dates. All sources on the Sony Hack report its begin date as _____, so write _____in Column G. Looking at the first report on p. 3, this incident had two events. The first hack did its work quickly, as when it was injected, it immediately went to work and stole 100 terabytes of files from Sony on that day. Therefore, it ended almost as quickly as it started. So we code the end date for the following day, on_____. However, the hackers came back on _____, where they injected wiper malware to erase all data on the networks' computers on _____. This is the end date for the Sony Hack.

The political fallout from the Sony hack lasted much longer than a day, as the hackers released the stolen files incrementally and the US government responses were also a series of incidents that lasted well over a month. The second article gives a detailed timeline of the hack. Yet we are only coding the time it took for the cyber incident to begin and do its malicious activity. The Sony Pictures' network was paralyzed for days, even weeks. Yet for the Sony Hack, it was less than a 24 hour period from the time it infiltrated to the time it was discovered and expunged from the Sony Pictures' network.

I. What type of interaction was the Sony Hack? According to the coding rules, the Sony Hack 1 is an offensive strike because its intent was to steal sensitive information from a multinational corporation. This was a clear case of an _____campaign, so we write ___ in Column I.

J. Wiper malware, which is usually a logic bomb method that erases data from networks, was used in the Sony Hack. What is unique about the Sony Hack is that wiper malware was used to steal the troves of information. How did the perpetrators get into the Sony network? Most reports are saying that it the physical location of Sony Pictures was very insecure, and that persons sympathetic to the GOP's cause were able to attain keycards, passwords, and other important information from the employees of Sony. There was therefore no Trojan Horse-style entry for the Sony hack. We will therefore code the hack as a logic bomb, or 4.1 in Column J.

K. Was it an Advanced Persistent Threat? Yes. The hackers went after a specific target and planned for months before infiltrating the network. The wiper malware was designed to steal and then wipe out the information on the hard drives of these computers, making it a somewhat tailored attack. Furthermore, the attackers warned Sony before the attack to not release *The Interview*, so secrecy was never a mode of operation for the North Korean hackers.

L. What type of target is Sony Pictures? We code this target as __, _____, because_____.

M. Who was the initiator? It is clear that _____ was the initiating state. We write its country code, _____, in Column M.

N. What was the objective of the initiator? Usually wipers, or logic bombs, are used to erase information. Not in the Sony Hack case. The infiltration was to wipe out all data only after a trove of information was stolen in the network. Therefore, we do not code it as ___, disruption, nor do we code it as ___, to coerce a state, as the movie was eventually released. The Sony Hacks' primary objective was to steal sensitive information and release it publicly, and we code it as ___ in Column N.

O. What was the political objective of the Sony Hack? _____
_____.

P. Did the objective work and change Sony's behavior? _____.

Q. Was there a 3rd party that helped North Korea initiate the Sony Hack? All reports on Sony report North Korea as the only initiator. We write ___ in Column O.

R. Did the Sony Hack target another country? No, it is apparent that Sony Pictures in the United States was the sole target. We write ___ in Column P.

S. How did North Korea respond to the Sony Hack when it was accused of launching it? All documents report that Iran denied its part in the incident. We therefore write ___ in Column Q.

T. How do we rank the SonyHack on the severity scale? The Sony Hack is clearly a ___ on the severity scale, it targeted a specific company and stole its critical information in a country. It did not merely deface users screen nor did it just shut down the system for a few hours. It stole and erased digital data which led to a response from the U.S. government, which eventually led to economic sanctions on North Korea. It was therefore too severe to be coded as ___. Furthermore, it did not steal information on a larger sector-wide scale, only Sony Pictures, so we do not code it as a ___. *The Interview* was also released on other forms other than theaters, so the intent of the attackers to stop its release also failed.
S. What type of damage did Sony Hack inflict? As soon as the intrusion entered Sony's network it began to steal data, and then proceeded to erase it. Its effects were felt immediately. It also did what its initiators wanted it to do, which was steal and erase data in Sony Picture's system. It succeeded. We therefore code the Sony Hack as a _____ and_____ incident, and write __ in Column S.

U. We copy and paste all sources we used to code Sony Hack in this column. We used 2 for this case.

V. In this column we write a short summary of Sony Hack. Here is the example:
"_____
_____
_____."

References:

Goodman, Will. 2010. "Cyber Deterrence: Tougher in Theory than in Practice?" *Strategic Studies Quarterly* Fall 2010: 102-135.

Jones, Daniel M., Stuart A. Bremer, and J. David Singer. 1996. "Militarized Interstate Disputes, 1816-1992: Rationale, Coding Rules, and Empirical Patterns." *Conflict Management and Peace Science,* 15 (2): 163-215.

Klein, James P., Gary Goertz, and Paul F, Diehl (2006) The new rivalry dataset: procedures and patterns. Journal of Peace Research 43 (3): 331-348.

Rid, Thomas and Ben Buchanan 2014. "Attributing Cyber Attacks."*Journal of Strategic Studies*, DOI: 10.1080/01402390.2014.977382

Thompson, William R. (2001) Identifying rivals and rivalries in world politics. International Studies Quarterly 45 (4): 557-86.

Valeriano, Brandon and Ryan C. Maness. 2014. "The Dynamics of Cyber Conflict between Rival Antagonists, 2001-2011." *Journal of Peace Research*, 51 (3): 347-360.

Valeriano, Brandon and Ryan C. Maness. 2015. *Cyber War versus Cyber Realities: Cyber Conflict in the International System* (New York: Oxford University Press).