Codebook for the Dyadic Cyber Incident and Dispute Data, Version 1

Brandon Valeriano
And
Ryan C. Maness

The Dynamics of Cyber Conflict between Rival Antagonists, 2001-2011
Journal of Peace Research
http://www.prio.no/jpr/datasets

Overview

This codebook presents a point of reference for variables for dyadic rival states that are in Valeriano and Maness's Cyber Conflict Data Project for the years 2001-2011. Rival dyads are extracted from the Klein, Diehl and Goertz (2006) enduring rival dataset as well as Thompson's (2001) strategic rival dataset. Each dyadic rival engaged in cyber conflict has two states involved, on opposite sides of the cyber incidents and disputes (see below for terminology). Each of these two states was directly involved in cyber incidents and disputes against the other, where the initiation must come from the government or there must be evidence that an incident or dispute was government sanctioned (see below for attribution 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, security companies, search engine corporations). The dataset does not include multi-targeted cyber incidents or disputes; these types of incidents or disputes are only coded at the dyadic level. Third parties are noted and coded separately. Both sides of the cyber incidents or disputes must be rivals; non-rival dyads are not included in the dataset. Furthermore, a dispute ends after there has been a substantial period of time before the next; or if the tactics by the initiator substantially changes. There are 126 rival dyads in the data, where this version only finds evidence of cyber conflict among 20 of these rival dyads for the time period of 2001-2011. 110 cyber incidents within 45 cyber disputes are present during this time period.

This is an active dataset and will be maintained, where more rival pairs of states as well as more cyber incidents and disputes is expected.

Specific Procedures

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

What to look for and record:

a. The rival dyad
b. Start and end date of interaction, cyber-incident and dispute
c. Type of interaction (nuisance, defensive, offensive) for both incidents and disputes (see different coding procedures)
d. Method of interaction/incident, 1-4 with decimal denotations for infiltrations (methods are

listed below) for both incidents and disputes (see different coding procedures)
e. The type of target (private/non-state, government non-military, government military)
f. The initiator of the interaction (use COW country codes, when it is a two-way operation, enter both country codes)
g. The foreign policy motive of the initiator for disputes only
h. Whether or not a third party was involved in the initiation (other state, rebel group, corporation) 1 = yes, 0 = no
i. Whether or not a third party was a target of the interaction 1= yes, 0 = no
j. Whether or not an official government statement was issued by the initiator, 0= no comment, 1= denial, 2= acceptance, 3-multiple
k. Damage level on the 1-5 scale level, given below for both incidents and disputes (same procedure for both)
l. The News Source for the cyber interaction
m. Any special notes pertaining to the interaction
n. Enter information in excel sheet for each rival dyad

Once these procedures are finished, attribution 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 disputes can be a problematic issue. 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 attribution 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. Attribution 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.

Methods of cyber-incidents
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.

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) Packet sniffers are software designed to capture information flowing across the web. 5) 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.

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 packet sniffer, .5 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**

Interaction type

1- Nuisance (probing, disruption, chaos)
2- Defensive operation (Cisco Raider, Buckshot Yankee, Israeli operations against cyber-jihad)
3- Offensive strike (Ghost Net, Shady RAT, Stuxnet)

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)

Severity (scale the same for incidents and disputes, for disputes, the highest incident is the severity code for the dispute)

1- Minimal damage (State dept website down, most defacements)
2- Targeted attack on critical infrastructure or military (financial sector blip, DoD hacked)
3- Dramatic effect on a country's specific strategy, (Stuxnet, Flame, jet plans, R & D)
4- Dramatic effect on entire country or large region (Power grid down, stock market collapse)
5- Devastating effect on country, (Catastrophe)

Coding for cyber disputes: 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.

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

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

Methods for disputes

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

Objectives for initiators

1- Disruption (take down websites, disrupt online activities)
2- Theft/Espionage (steal sensitive information or strategies)
3- Change in behavior (abandon nuclear program, withdraw troops)

Variables in Valeriano_Maness_JPR_CyberConflict_DataProject_v1.csv

Variable Number	Variable Name	Variable Description
1	Cyberdisputenum	Cyber dispute number
2	Cyberincidentnum	Cyber incident number
3	rivnumb	Rivalry number

4	Rivalry	Rivalry (COW codes)
5	Rivala	First rival in dyad
6	Rivalb	Second rival in dyad
7	Beginr	Rival start date
8	Endr	Rival end date
9	totMID	Total militarized interstate disputes between rivals
10	rivtyp2	Rivalry? y/n
11	rivtypold	Rivalry type? Enduring, Proto, Strategic
12	Name	Name of cyber incident or dispute
13	interactionstartdate	Cyber incident or dispute start date
14	interactionenddate	Cyber incident or dispute end date
15	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)
16	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 - Packet sniffer 4.5 – Keystroke logging 5- Vandalism and Denial of Service (disputes only) 6- Intrusion and Infiltration
17	APT	Advanced Persistent Threat? 1- Yes, 0- No
18	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)
19	initiator	State that initiated the incident or dispute (COW code)
20	initiator objective	Objective of the initiating state (disputes only) 1- Disruption 2- Theft/Espionage 3- Change Target Behavior
21	3rdpartyinitiator	Third party involved with initiating state? 1- Yes, 0- No
22	3rdparty target	Third party involved as a target? 1- Yes, 0- No
23	Govtstatement	Statement from the initiating state? 0- No comment, 1- Denial, 2- Acceptance, 3- Multiple statements

24	Severity	Severity level of incident or dispute (for disputes code the highest incident severity 1- Minimal damage 2- Targeted attack on critical infrastructure or military 3- Dramatic effect on a country's specific strategy 4- Dramatic effect on entire country 5- Devastating effect on a country
25	Source	The news source for the cyber interaction
26	Notes	Any special notes pertaining to the interaction

References:

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.

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