The underlying causal network from global dyadic events: allies and rivals in international relations

Lei Jiang
Microsoft
Bellevue, WA, USA
lionelchange@hotmail.com

***Abstract*—The ever-increasing amount of web information usually leads to more discovery of implicit relations using data-driven methods. While country-country relations are not exactly new for human learning, we illustrate that one can extract the knowledge of such sophisticated relations solely from data. Using the global dyadic events data that record the involving parties and timestamp, we identify the underlying causal network with common initiators or reactors, as well as the nature of pairwise relations as a counterfactual reasoning problem, thereby deriving the significance of causal effects in the network. To formalize the discovery process, we first validate some country-country relations in a small-world network structure after exploitative runs over the entire data set, suggesting the traditional allies and rivals as we know in the real world. Further, we use sentimental analysis based on Wikipedia text data to match the significance of bilateral relations in the causal discovery. Our results show that using time series data of dyadic event counts alone, combined with lexical-based validation, can extract the knowledge of bilateral and multilateral relations and has potential to be a means to measuring and monitoring the climate of complex geopolitics.**

***Index Terms*—causal network, dyadic events, GDELT, counterfactual, bilateral relation**

I. Introduction

Country-country relationship, usually defined by bilateral or multilateral relations, is known to be sophisticated. Normally, people can learn such relations in the world through traditional media reporting and history books as a human cognitive process: diplomatic events between countries, cultural and religious conflicts, allies and rivals in past wars, and etc. While official news reports would rarely spell out the extent of a non-ally relation, it could be judged from historical events, and also the frequency of reported interactions, and the tone of them. A link between two nations is easy to draw but the nature of it could be complicated: two countries may have close ties between non-official organizations but call each other "competitor" or "rival", and one country could be a mediator with regard to the matters between two other countries.

In the era of web and social media, the increasing online presence of political figures and national and regional entities[1] is making the digitization of diplomacy a reality. It provides abundant new resources and tools, under a setting of social networks, for discovering the underlying characteristics and principles: e.g. the usage pattern of Twitter and its contents are analyzed for certain relations [1]–[3], the network structure based on follower-following of Twitter accounts of embassies is explored [4]. Considering the fact that countries are actually not equal in economy power and level of participation in international affairs, an undifferentiated social following, like in the real-world people-people social network, demonstrates "weak-tie" friendship. However, it may not capture the significance and doesn't reflect the real role that a country plays. In this paper, the author proposes a new methodology based on causality to discover bilateral and multilateral relations.

Causal discovery has been an active topic in statistics and machine learning for many years. Inspired by its wide range of applications in econometrics [5], [6], biomedical [7], [8], social sciences [9], [10] and etc., many approaches emerge as they seek to find the links among multiple variables. Among these, depending on the specific scenario, causal inference can be made *i)* through a pair-wise statistical significance test or *ii)* by searching through a network with subtle relationships in between: e.g. a single link between corporate philanthropy and revenues is up for testing with the null hypothesis that "more giving will result in more customer satisfaction, and thus bringing more future sales".

Meanwhile, discovering causal effects presents several technical challenges in the context of international relations:

- Causal effects are multi-faceted in bilateral and multilateral relations.
- Low data quality or skewed distribution could result in bias and even lead to false causal relationship.
- Differentiating correlation with causality requires additional steps of validation.

We address these questions in a formal framework that utilizes the Global Database of Events, Location and Tone (GDELT) as well as crawling related Wikipedia web data for validation. First, conceptually, we consider that a *causal link* can be established between two different geopolitical entities over another entity (e.g. the action taken by Russian on Syria would cause United States' reaction). Accordingly, we can extract structured dyadic event counts from GDELT data sets and use PC algorithm [11] as a baseline to exploit all potential pairwise causal links as well as DAGs (directed acyclic graph) and we also evaluate the significance of bilateral relations as a counterfactual problem. Then, having the discovered structures as a set of hypothesis, we further adopt an automated text-

[1]in this paper, we interchangeably use country, nation, and entity for the same subject.

based approach to perform the validation process: we scrape Wikipedia pages on bilateral relations to match all the words with dictionaries of sentiment/effect and the frequency-based score becomes a quantifiable measurement for checking each causal link. While the latter step of validation is considered a more reliable approach as it is based on documented words that present a ground truth, a set of hypotheses are tested and verified with regard to the causal links found . This consistency shows the validity in detecting causal relationships in multilateral relations from the dyadic event counts (as a structural method).

II. Related Work

Causality, or causation, indicates a deep and intrinsic relationship between certain entities. More than easy-to-calculate correlation, it takes conditional probability into account and aims to remove spurious links brought by confounding or latent factors [12]. In a canonical framework, the PC algorithm [11], as a constraint-based learning algorithm, performs statistical tests to derive a set of conditional independence and dependence statements: the skeleton, which is an undirected graph, is found by pairwise comparisons and then directions are added to graph with collision removed to ensure its acyclic nature, resulting in a Bayesian belief network. This approach has been extended for high-dimensional space [13] and nonparanormal models [14].

While researchers in political science have long argued for the existence of causation in international relations [15], [16], to the author's best knowledge, there is no prior work that presents quantitative and data-driven approaches to explicitly deriving causality between geopolitical entities. In this study, countries and regions in the world are considered as a social network and the causal relationship between them can be interpreted as peer effects. In economics and statistics, this is a familiar subject [17], [18]. Peer effects can be formulated as a type of causal inference, and the treatment effects can thereby measured through randomization design, where a linear model is trained using parametric methods. In data mining related applications on the global dyadic events data, [19] also uses GDELT data set to model Bayesian Poisson tensors for predictive analysis as a regression problem, [20] predicts future level of violence in certain regions and [21] constructs a Bayesian network on various types of actions from countries. In this paper, we put an emphasis on finding causal links and exemplifying the relationship between geopolitical entities hidden in the data, while the causal links can actually help with regression as well.

Instead of establishing a theoretical framework, our proposed approach aims at validating the graph structure as well as using another data source. And we treat these two steps as "non-text-based" and "text-based" respectively. Text mining and sentimental analysis has been an active area in data mining: when relations to be extracted from comparative sentences, where entities can be evaluated for rule derivation [22], a classifier works on the keywords and the categorization like non-equal gradable, equative and superlative results in features; while more entities can be involved in the context in social media text, they can be discovered and assigned by part-of-speech tagging, identifying opinion indicators and pattern matching [23].

TABLE I
Number of events ranked by country (initiator and reactor respectively) (1998-2007)

	Country	#Initiations	Country	#Reactions
1	United States	1157368	United States	1115651
2	Russia	573606	Russia	527904
3	Israel	485240	Iraq	502551
4	United Kingdom	371077	Israel	476792
5	China	367154	Palestine	363218
6	Iraq	343047	China	355157
7	France	289072	United Kingdom	341797
8	Japan	285731	Iran	272851
9	Iran	281072	Japan	263255
10	Palestine	279630	European Union	251667
11	Egypt	219457	France	250137
12	European Union	219225	Turkey	197102

III. Data Sets and Concepts

The Global database on events, location and tone (GDELT) [24] is a publicly accessible data set[2] that monitors the worldwide media sources in multiple forms (including print, broadcast and web), multiple languages and records context information of them. Also, it labels each such event with a standard format that encodes two involved entities (*initiator* and *reactor*), location, type and severity for each event record using CAMEO (Conflict and Mediation Event Observations) [25]. That is, the GDELT data is a complete time-dependent representation of the news around the world. As the event definition here is already directional with exactly one single initiator and one reactor, it is consistently applied in this paper throughout all the steps of data processing and causal reasoning.

While the raw GDELT data has a rich set of features, we aggregate the dyadic event counts between all the dyads, such as Japan $\rightarrow$ South Korea, rather than a country name alone. To map the event frequencies to real-world relations, the underlined assumption is that dyadic event frequency does reflect the intensity of diplomatic activities and its trend over time does indicate the development of bilateral relations. The ranking of countries by the number of initiation and reaction events are listed in Table I. Intuitively, the larger the number is, the more actively this geopolitical entity behaves. The distribution doesn't necessarily represent the wealth of a country but show the degree of becoming a spotlight as in media reports and discussions.

Figure 1-top gives a glimpse of the mutual relation by plotting event frequencies between China and United States, each of which as initiator respectively: as the blue and red curves highly overlap with each other, a strong correlation is easily seen; However, United States-China and China-United States events share a high correlation coefficient (Pearson's r)

[2]http://gdeltproject.org/

TABLE II
PEARSON'S R FOR US-OUTPUT EVENT FREQUENCY TO THE FIVE COUNTRIES

	Brazil	China	S.Korea	Japan	Russia
Brazil	1.00	-0.0352	-0.0488	0.0346	0.0377
China	-0.0352	1.00	0.128	0.0741	0.0826
Korea	-0.0488	0.128	1.00	0.208	0.134
Japan	0.0346	0.0741	0.208	1.00	0.102
Russia	0.0377	0.0826	0.134	0.102	1.00

0.928, but United States-Japan and United States-South Korea can only achieve 0.208. In sense of correlation, a symmetric matrix can thereby derived for the frequency of events from United States to each country in Table II (using data of 2000), where quite a few of them is less than 0.135.

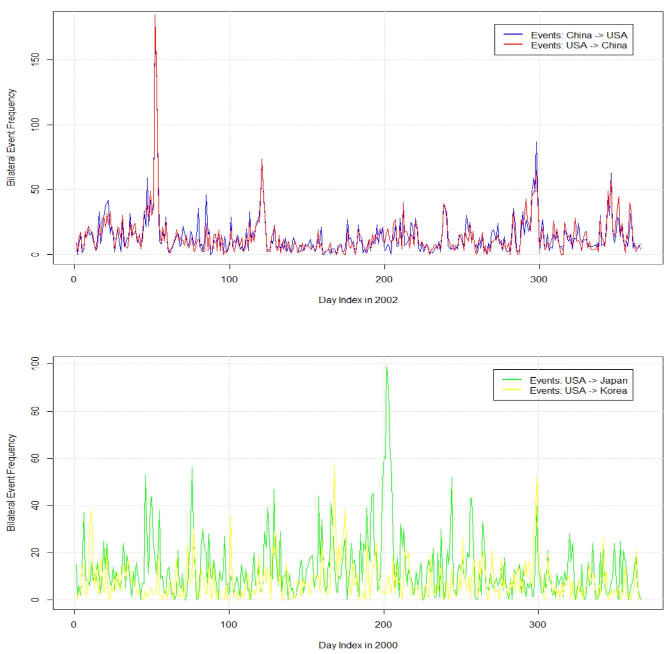

Fig. 1. top: The frequency of events between China and United States (each as initiator) in 2002; bottom: Events acted by United States to Japan and South Korea respectively in 2000

IV. CAUSAL NETWORK DISCOVERY

Dyadic events match with the multi-faceted nature of country-country relationship quite well: we may describe the relationship in a way like: Country A's diplomatic policy has an effect on how Country B is treating Country C. That is, we can first single out Country A, and its effect on all other countries can form a causal network (e.g. Country B would have a link to Country C, where Country A is, in fact, a confounder). The presumption is that Country A could be a confounder on all other countries' relation to each other (where it does exist in the real world for some powerful countries), and then we can derive a causal network for all other countries under A's influence. Not every country is necessarily affected by A, so that will be examined by cross-checking each of the causal networks with a different influencer in Subsection IV-B.

Thus, as we start from these causal networks under one country's influence, the following definitions are derived:

Definition 1 *Co-initiation link: when two dyads B and C show a causal link from B to C over a common reactor A from data sets* $(B \triangleright A)$ *and* $(C \triangleright A)$*, there is a co-initiation link from B to C associated with A, denoted as* $(B \rightarrow C \mid\triangleright A)$*.*

Definition 2 *Co-reaction link: when two dyads B and C show a causal link from B to C over a common initiator A from data sets* $(A \triangleright B)$ *and* $(A \triangleright C)$*, there is a co-initiation link from B to C associated with A, denoted as* $(B \rightarrow C \mid\triangleleft A)$*.*

The data used for each run of causal link discovery is a subset extracted from the entire set in order to avoid the spuriosity brought by data sparsity: a target entity is selected as initiator or reactor, and the other entities' events with this target are counted. Considering the entire 211-country data set is highly sparse, only top 19 geopolitical entities are selected for much of the causal discovery process: the United States (USA) , European Union (EU), United Kingdom (GBR), Russia (RUS), China (CHN), Japan (JPN), South Korea (KOR), North Korea (PRK), Germany (DEU), France (FRA), Pakistan (PAK), Afghanistan (AFG), Turkey (TUR), Iran (IRN), Australia (AUS), Canada (CAN), Iraq (IRQ), Israel (ISR) and Palestine (PSE).

In the following sections, some interpretations are provided based on public geopolitical knowledge, for the sole purpose of demonstrating that causal links have a real-world touch and showing the correctness of this methodology, even though it might not perfectly explain each and every individual causal link.

A. Co-initiation and co-reaction causal networks

The PC algorithm [11] is commonly used for causal inference, with an advantage of no hidden settings and no need to select variables. The steps of PC algorithm include: first constructing a complete directed graph and then test conditional independence for all the pairs to build a skeleton with all undirected links; then it identifies colliders and handles them with structural constraint (d-separation), which ensures no cycle would occur throughout the process; next, the directions can be set up and the remaining undirected links can be randomly assigned as long as it's a DAG.

Choosing a country one at a time as initiator and reactor roles respectively, the outcome of causal discovery comes from the input data of corresponding monthly-aggregated dyadic event counts. Figure 2 shows the DAG[3] with United States as target reactor (we use $\alpha = 0.05$ in conditional independence test).

Considering how the relationship on the graph could imply the real-world situation, in Figure 2-bottom, United States is

[3]In Figure 2, 3 and 4, some edges are bidirectional and represent undecided orientations. In a strict DAG, they would have a random orientation to ensure acyclic.

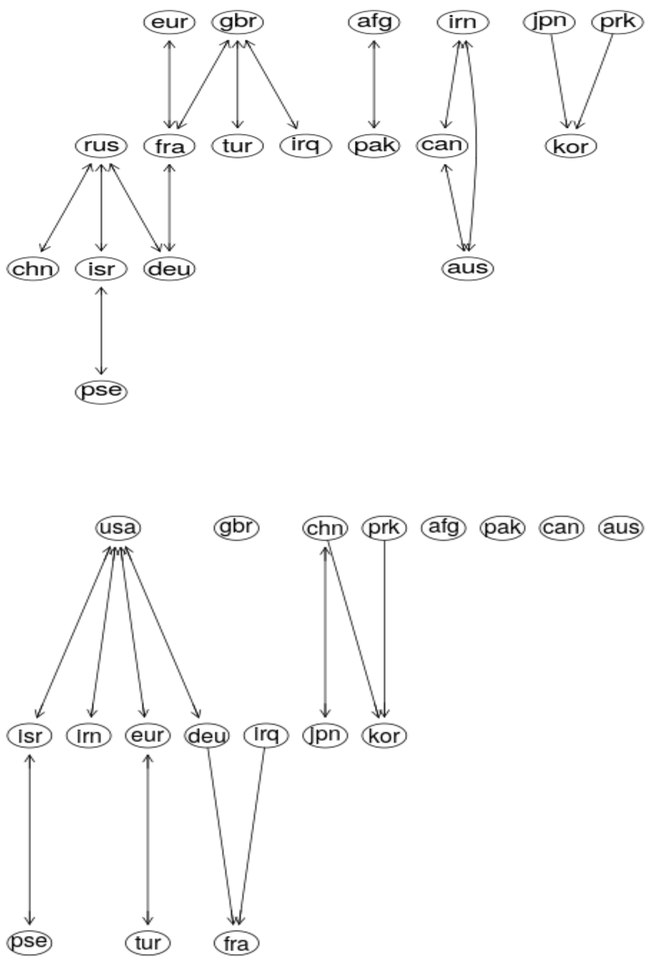

Fig. 2. Co-initiation causal networks with United States (top) and Russia (bottom) as reactor respectively using 1998-2007 data for other entities

clearly in the center as this node's outdegree is 4 (the largest). Also, it has an inlink from Iran, which can be interpreted as "the US policy to Russia depends on Iran's actions". In other words, "Russia could use Iran to influence its relation with United States", where such knowledge can be found in public online materials[4].

A quick integrity check is performed: switching from co-initiation to co-reaction networks, would all the arrows simply be flipped and the structure remain? Figure 3 illustrates such a co-reaction structure for Russia (comparable to Figure 2-bottom). While one exception is that the United Kingdom is singled out in co-initiation links while it gets involved in the co-reaction, where a plain explanation is that "UK doesn't necessarily follow what the US do with Russia, but when Russia made some deals with Iran, it would react". The same reasoning applies for (Australia $\rightarrow$ Afghanistan — $\triangleleft$ Russia).

[4]https://en.wikipedia.org/wiki/Nuclear_program_of_Iran

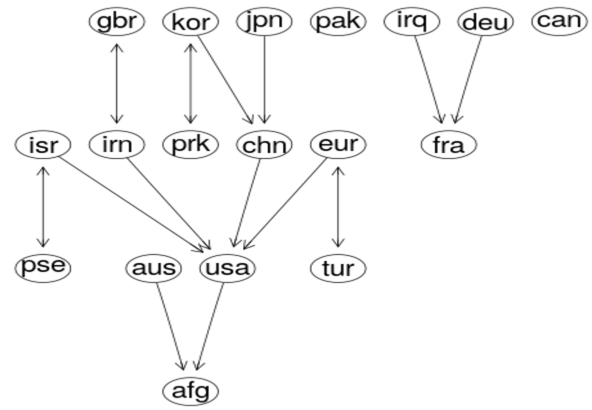

Fig. 3. Co-reaction causal network with Russia as initiator using 1998-2007 data for other entities

Diving deeper into the links, some exemplify geographic adjacency between two entities: e.g. South Korea is usually linked to North Korea or Japan; Israel is always connected with Palestine. While they are true in the real-world setting, it shows that using co-initiation and co-reaction causal networks we can answer questions like "*how other countries would be influenced with one's action*" where zero prior knowledge is given. A larger-scale exploitative run of PC algorithm over the entire data set is described in *Algorithm 1*.

B. Revisiting bilateral relations in a counterfactual view

With co-initial and co-reaction links creating an additional dimension, now we can give a closer review on a certain pair of countries, getting back to answering the basic questions like "how important the bilateral relation between Country A and B is".

Counterfactual reasoning has been used in several machine learning applications [26], [27]. The concept exists in assuming a "what-if" situation for the past events and derive offline experiments or simulations to find optimization solutions. Under this paper's framework, we can put counterfactual reasoning in this way: in a co-initiation causal network (under the influence of Country A, and this is the same for co-reaction networks), if we remove all the dyadic events between Country $B \rightarrow A$ and then re-construct this network, we might observe two *extreme* circumstances:

- the network structure changes a lot, OR
- the network structure doesn't change at all with the removal of B.

In the prior case, it suggests the bilateral relation between Country A and B is so important that it would reshuffle the network structure; and the latter case would mean this relation is not significant at all.

The following metric is then derived for quantifying the significance in terms of network structure difference:

Definition 3 *Significance of a bilateral relation*: in a co-initiation (or co-reaction) network under A, the significance of relation B $\rightarrow$ A is represented by the net change of distinct edges in the re-constructed network after removing all the events B $\rightarrow$ A (A $\rightarrow$ B for co-reaction networks).

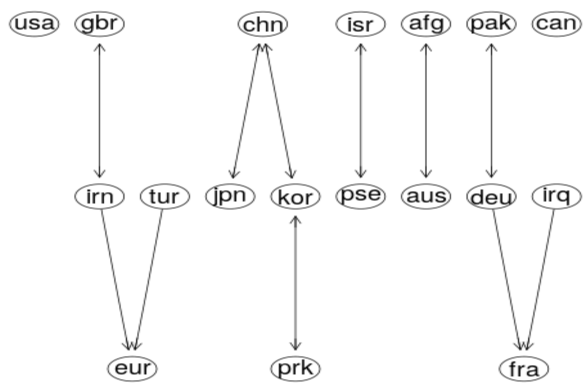

Fig. 4. Co-reaction causal network under a counterfactual scenario with Russia as initiator and United States events removed (1998-2007)

Figure 4 illustrates the counterfactual case with United States singled out in a co-reaction network under Russia (comparable to Figure 3). Obviously, it turns into 7 disconnected components (from 4). Some cliques in the original network, such as Israel-Palestine, do remain, but now they are separated from others. *Algorithm 2* in Section illustrates a full-range run of counterfactual reasoning on the entire data set.

Here we use the *net change* between two networks rather than a normalized one. Our experiment shows the result from net change looks more evenly distributed (as in Table III), as the top relations involve many countries. Instead, it skews more towards the presence of the United States after normalization.

In the next section, we will validate such causal links from network structure as well as from external text resources.

V. Validating the Causal Links

The validation consists of two steps: first, a run-through over all the countries as targets provides an expanded and more complex picture for examining the nature of the causal links found; then we use text mining method to scrape related ***Wikipedia*** web pages as an independent source, where the links are scored by sentiment lexicon matching.

A. Exploit causal modeling over the entire data set

The validation stage starts from further identifying causal links from the discovery process, attempting to get rid of potential spurious effects in a single run. As each time we run a causal discovery routine with one single target (as initiator or reactor) and the results are to be combined. With the combination, it also expands the relationship map to more entities: e.g. if we see both (B $\rightarrow$ C $|\triangleright$ A) and (B $\rightarrow$ C $|\triangleright$ D), the causal link $B \rightarrow C$ appears stronger.

Algorithm 1: Constructing co-initiation and co-reaction causal networks

Input: all the pairwise time series where data a point represents the event counts at (C_i, C_j, t_k), $i, j \in 1...n$ and $i \neq j$. It is different from (C_j, C_i, t_k) when entity C_i and C_j exchanges as initiator and reactor.

Begin

for each entity C_i $(i \in 1...n)$
$--$ run a causal discovery routine with the subset of data represented by $(C_i, C_{1...i-1,i+1...n}, t_{\text{inf}})$
$--$ run a causal discovery routine with the subset of data represented by $(C_{1...i-1,i+1...n}, C_i, t_{\text{inf}})$

end for

Combine the above results by same causal link over varied targets, and remove links such that $(C_i \rightarrow C_j| \triangleright C_p)$ exists but $(C_j \rightarrow C_i| \triangleleft C_p)$ doesn't hold.

End

Output: All the causal links denoted as

$$(C_i \rightarrow C_j| \triangleright C_p 1...C_p r)$$

$$(C_i \rightarrow C_j| \triangleleft C_q 1...C_q s)$$

We define *special causal link* and based on its appearance pattern after the initial discovery process.

Definition 4 *Special causal link* is a co-initiation or co-reaction link that shows some irregularity in appearing in the graph with different targets or a target being another role: *a.* the link reverses its orientation when the target switches from initiator to reactor (and vice versa); and *b.* the link exists/doesn't exist only if the target is a certain entity.

Definition 5 *Stable causal link* is a co-initiation or co-reaction link that shows a pattern in appearing in the graph with different targets or a target being another role: *a.* the link reverses its orientation when the target switches from initiator to reactor (and vice versa); and *b.* the link persists under two or more than two targets.

This definition further probes the uni-directional attribute in the network, while eliminating some bias or numerical error by requiring it to show up in both initiator and reactor roles. Thus, after an exploitation of running PC algorithm through all the important entities being either role (set α lower to 0.005 for raising the bar of causal link). From the findings, interpretations can be made on some interesting links: the special causal link (EUR $\rightarrow$ USA $|\triangleright$ RUS), meaning the European Union's actions on Russia would make the United States to follow and Russia is the only possible target for this link; (AFG $\rightarrow$ PAK) appears as a stable link for a set of entities as the target reactor: IRN, USA, JPN, GBR, DEU, FRA and EUR, and this could be related to the military actions in Afghanistan after 9/11.

Out of the total 598 links, there are 72 special causal links and 10 stable ones discovered. While the distribution follows a power law, those links with more than 1 common target indicates peer effects and the prevalence of special causal relationship, with reasonable interpretations on many cases, exemplifies the complication in world politics. In particular, 36 out of 72 involves the United States, Russia or China as a part of the causal link, while another 13 can be added with any of the three serving as a target entity. Using this approach, the influence of titans could be quantitatively measured.

B. Ranking the significance of bilateral relations

With all the co-initiation and co-reaction networks as input, counterfactual reasoning can help reduce the dimension of common initiator/reactor and convert it back to bilateral relations, as stated in Section IV-B.

Algorithm 2: Ranking the significance of bilateral relations

Input: all the coinitiation and co-reaction networks constructed in Algorithm 1

Begin

for each causal network where C_m is the co-reactor/co-initiator:
$--$ for each entity C_i ($i \in 1...n$, $i \neq m$):
$----$ set all such $(C_i, C_m, t_{inf}) = 0$
$----$ run a causal discovery routine with the subset of data represented by $(C_i, C_{1...i-1,i+1...n}, t_{\inf})$
$----$ compare this new network to the original one and get triples $(C_i, C_m, E_d iff)$
$--$end for
end for

Aggregate all such $(C_i, C_m, E_d iff)$ and merge with $(C_m, C_i, E_d iff)$ by summing up $E_d iff$ over each C_m.

Rank them based on aggregated $E_d iff$.
End

Output: Ranked bilateral relations denoted as (C_r, C_s), where $r, s \in 1...n, r \neq s$.

The significance of a bilateral relation actually shows the power of influencing other countries under this causal modeling framework, instead of its real-world political effect. In the top relations discovered, the Afghanistan and Iraq wars, which happened during the period, have factored in as allies followed suit. Some notable local conflicts, such as Israel-Palestine (ranked 31 out of 171) and the Korean Peninsula (KOR-PRK, ranked 114), don't get a high note. In part, the conflicts are relatively contained rather than more entities getting involved. In other words, a country's stand on them is consistent, yet no substantial involvement is needed.

TABLE III
MOST SIGNIFICANT BILATERAL RELATIONS FROM CAUSAL DISCOVERY OVER 1998-2007 DATA

Rank	Relation	Rank	Relation	Rank	Relation
1	RUS-USA	6	AUS-USA	11	FRA-ISR
2	AFG-USA	7	CAN-GBR	12	EUR-IRQ
3	AFG-PAK	8	PAK-USA	13	FRA-USA
4	IRQ-USA	9	GBR-PAK	14	FRA-GBR
5	EUR-USA	10	AUS-CHN	15	CHN-GBR

C. Sentimental analysis for validation

With the interest in mind in trying to systematically find the ground truth of causal links between nations, text mining against an independent text data source is easily conceivable: while causal links are established through numeric data (dyadic event counts) without any priors, the texts from news are the real clue from text sentiments. So, we consider ***Wikipedia*** as a reliable web text source as it has created (through editors) and maintained the pages that describe major bilateral relations in the world.

By retrieving such a Wikipedia page, we use a sense-level lexicon dictionary to measure the sentiment [28]. Matching related words to the content makes the level of effect quantifiable. In this scenario, we specifically probe the comparison in a triangle relationship: e.g. when there's a co-initiation link (B $\rightarrow$ C $|\triangleright$ A), we would evaluate the contents of (B $\triangleright$ A) and (C $\triangleright$ A). Given that these are two distinct relations and independent sources, it's too idealistic to say that every piece of the graph structure of causal links can match the text-based approach. However, we do see maintained consistency in key features as we switch from event counts and causal graphs into texts.

In applying the lexicon matching, the positive and negative effects can be scored respectively and thereby metrics can be derived as *text significance score*, E_{sig}, *positivity score*, E_{pos} and *effect score*, E_{effect} using the following equations.

$$E_{sig} = C_{pos} + C_{neg}$$

$$E_{pos} = C_{pos}/(C_{pos} + C_{neg})$$

$$E_{effect} = (C_{pos} + C_{neg})/C_{total}$$

where C_{pos} and C_{neg} are the counts of positive or negative effect lexicons; C_{total} is the total word count on the web page.

The hypotheses from the graph structure of causal links can be proposed and tested:

Hypothesis 1: The bilateral relations involved in the causal links are more significant than average.
After performing a scoring procedure that takes into account all words having a sentiment effect, the average significance score is 10% higher than that of the whole set of pairs available from Wikipedia. Also, as Wikipedia doesn't have all the pairwise bilateral relations pages, the availability of such pages is 86.5%. Singing out all the bilateral relations that we find any causal link in the co-initiation or co-reaction networks, the level of absence slightly reduces: the availability is now 90.0% out of the 130 relations extracted from causal links. Note that we consider two pairwise relations "United States and Japan" and "Japan and South Korea" as *involved* in a causal link (USA $\rightarrow$ JPN $|\triangleright$ KOR).

Hypothesis 2: Measured by *effect score*, the bilateral relation from and within causal links represent more ties, connections or historical activities between the two entities.

The *effect score* calculates the proportion of "words/lexicons with an effect" out of the entire text. The higher the score is, it shows a more storied past and present in the bilateral relation. It doesn't simply represent how long the history is as it's a normalized score. In our result, "involved " relations stand for 1% higher effect score than the average, while relations indicated by causal link achieve a 3.5% high.

Hypothesis 3: Stable causal links have a larger significance score as well as a higher effect score.

The stable causal links, as they do represent an even stronger geopolitical connection (e.g. Iran $\rightarrow$ United States, but not the other way around), stand out from the metrics in both ways. With the calibration mechanism in place that those non-reversing causal links after an initiator-reactor exchange are discarded in verification, it would make stable links less error prone. That doesn't mean special causal links are less useful. Instead, special causal links, or even those we temporarily eliminate for non-reversing, may reflect more uni-directional nature in the global situation.

Meanwhile, the *positivity* score doesn't reflect an obvious change between different groups of relations. It may suggest that the Wikipedia articles are monitored by editors and they are meant to be a neutral information source.

After going through this multi-step validation process, the causal links turn out to represent a relatively deeper level of relations than those unconnected entities. Coherence is seen between the non-text (structural) and text based approaches, and the hypotheses examined indicate a good likelihood that many causal links found are true.

VI. Discussion and Analysis

While existing work in causal discovery on observational data focuses more on "factors", where the two ends of a causal link are not sovereign entities, we first identify the causality between two dyads (co-initiation and co-reaction networks) and then return to the significance of bilateral relation after an exploitative process. That is because the event counts data (time series) present itself in such a form that there must be an initiator and a reactor in every single event. Throughout the paper, the causal discovery process doesn't pose any prior knowledge on top of observational data. Moreover, it starts from assuming every entity can be a confounder, so an implicit filtering embeds in Algorithm 1 and 2 via exploitation and ranking.

Trying to further interpreting the nature of the causal links, it would be too arbitrary to say that the one entity has a direct effect to the other in the same causal link. In the real-world, it is conceivable that one region could have an *indirect effect* for its stand on some issues, and such a stand can influence other countries which could exert the impact directly. In causality, mediators may well exist [29], and to distinguish a mediator from a true confounder, granular data more than dyadic event counts is in need.

Data quality plays a role throughout the causal discovery process. We choose the period 1998-2007 for the following reasons: *a.* the event counts used for causal discovery are dictated by news volume and a burst of news volume due to the emergence of new media would actually produce noise; *b.* 10 years is appropriate for causal reasoning when it aims for stable relationship. In fact, country-country relations doesn't stand still in a longer period and new events and dynamics in the world are gradually changing them. If we want to capture such dynamics, methodology and algorithm aside, it would require much more data cleaning and preprocessing to counteract the noise that finer granularity of data introduces as a byproduct, in order to achieve this ambitious goal. Meanwhile, the PC algorithm, as used in our work, is established on this statistical foundation and consists of steps in removing confounding effect from graph, and thereby doesn't have a time dimension in its modeling. Practically, in terms of parameter tuning, we do see that a high level of significance (e.g. $\alpha = 0.005$) and a short data span like 2 years would lead to very few causal links found by PC algorithm.

On the other hand, Granger causality [30] uses short-term regression in its core and is still active in some application areas, where tremendous time complexity makes learning Bayesian networks infeasible. Meanwhile, it is important to validate the potential causal links found in order to avoid spurious causal effects, and in some cases, differentiating correlation from causation, as well as identifying instantaneous causality, would play a role in the analytical results. Granger causality is native to time series and arguably more efficient for such data. Echoing to the existing research where [12], [31] point out Granger causality's limitation in failing to find confounders, we target causal links that are "as stable as possible" and thereby don't include results with Granger causality, in spite of that having generated a denser causal network in our experiments as expected [32].

In the validation phase, we systematically run the causal discovery process as exploitation and use text-based methods to provide reference scores. We propose several hypotheses with regard to the significance of causal links, and then we derive content-based scores to provide quantifiable measurements. Especially, Hypotheses 1-2 point to the overall significance and Hypothesis 3 is a deeper indication about graph structure. Wikipedia pages on bilateral relations is a good public data set that records the important diplomatic activities and characterize the relation using featured words. It is worth reiterating that country-country relationship is sophisticated in nature, so getting a labeled data set as ground truth, in a pure black-or-white style, is difficult and would become a sensitive matter.

Digging into real-world bilateral and multilateral relations is multidisciplinary in essence. We hope that the results of this paper[5] could serve as a baseline for a large-scale system with fresh data.

VII. Conclusion

In this paper, we tackle "causal discovery" in the realm of international relations as a web data mining problem: by aggregation dyadic events between pairwise geopolitical entities, we apply a formal framework in the discovery process: *i.* Using PC algorithm to identify causal links where we have shown are stable; *ii.* Applying a counterfactual reasoning process to evaluate the significance of bilateral relations and *iii.* Applying validation using both structural and text-based methods. The causal links found, through validation, are in accordance with the current world order and largely match what human can infer from news: PC algorithm's capability in locating stable links fits with this scenario, and the exploitative run of systematic causal discovery, including counterfactual reasoning, does extract valuable causal links out of the entire problem space. Our overall methodology and results demonstrate the application of causal modeling towards sophisticated relations in this interesting domain.

References

[1] C. Duncombe, "Twitter and transformative diplomacy: social media and Iran-US relations,", *International Affairs*, 2017.

[2] M. Ittefaq, "Digital diplomacy via social networks: a cross-national analysis of governmental usage of Facebook and Twitter for digital engagement," *Journal of Contemporary Eastern Asia*, vol. 18, No. 1, pp. 49-69, 2019.

[3] S. Shahin and Q.E. Huang, "Friend, ally, or rival? Twitter diplomacy as 'technosocial' performance of national identity," *International Journal of Communication*, vol. 13, pp. 5100-5118, 2019.

[4] E. Sevin and I. Manor, "From embassy ties to Twitter links: comparing offline and online diplomatic networks," *Policy and Internet*, 2019.

[5] Cheng, B.S. and Lai, T.W., "An investigation of co-integration and causality between energy consumption and economic activity in Taiwan," *Energy Economics*, Vol. 19(4), pp. 435-444, 1997.

[6] Asimakopoulos, I., Ayling, D. and Mahmood, W.M., "Non-linear Granger causality in the currency futures return," *Economics Letters*, Vol. 68, pp. 25-30, 2000.

[7] Tam, G.H., Chang, C., Hung, Y.S., "Gene regulatory network discovery using pairwise Granger causality," *IET System Biology*, Vol. 7(5), pp. 195-204, 2013.

[8] Kleinberg, S., Hripcsak, G., "A review of causal inference for biomedical informatics," *Journal of Biomedical Informatics*, Vol. 44(6), pp. 1102-1112, 2011.

[9] Reiss, J., "Causation in the social sciences: evidence, inference and purpose," *Philosophy of the Social Sciences*, Vol. 39(1), pp. 20-40, 2009.

[10] Lev, B., Petrovits, C. and Radhakrishnan, S., "Is doing good good for you? How corporate charitable contributions enhance revenue growth," *Strategic Management Journal*, Vol. 31, pp. 182-200, 2010.

[11] Spirtes, P, Glymour, C, Scheines, R. Causation, Prediction, and Search. Adaptive Computation and Machine Learning, 2nd edition. MIT Press, Cambridge, 2000.

[12] Spirtes, P. and Zhang, K., Causal discovery and inference: concepts and recent methodological advances, *Applied Informatics*, Vol. 3(3), 2016.

[13] Kalisch, M. and Buhlmann, P., Estimating high-dimensional directed acyclic graphs with the PC-algorithm, *Journal of Machine Learning Research*, Vol. 8, 2007.

[14] Harris, N. and Drton, M., PC Algorithm for nonparanormal graphical models, *Journal of Machine Learning Research*, Vol. 14, 2013.

[15] Kurki, Milja. *Causation in International Relations: Reclaiming Causal Analysis*. Cambridge University Press, 2008.

[16] Robinson, James , What causes IR? Differentiating causal analyses in international relations, APSA Annual Meeting Paper, 2011.

[17] Goldsmith-Pinkham, P. and Imbens, Social networks and the identification of peer effects, *Journal of Business and Economic Statistics*, Vol. 31, 2013.

[18] Toulis, P. and Kao, E., Estimation of causal peer influence effects, *Proc. of Int'l Conf. of Machine Learning*, 2013.

[19] Schein, A., Paisley, J., Blei, D.M., Wallach, H., Bayesian Poisson tensor factorization for inferring multilateral relations from sparse dyadic event counts, *Proc. of ACM KDD Int'l Conf. of Data Mining and Knowledge Discovery*, 2015.

[20] J. E. Yonamine, Predicting future levels of violence in Afghanistan districts using GDELT, working paper, 2013.

[21] Jalal-Kamali, Ali and Pynadath, David, Toward a Bayesian network model of events in international relations, in *Proc. of International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation* (SBP-BRiMS), 2016.

[22] Jindal, N. and Liu, B., Mining comparative sentences and relations, *Proc. of AAAI Conference*, 2006.

[23] Ding, M., Chen, Y., and Bressler, S.L. *Granger causality: Basic theory and application to neuroscience*. In Schelter. S., Winterhalder, N., & Timmer, J. Handbook of Time Series Analysis. Wiley, Wienheim, 2006.

[24] Leetaru, K., Schrodt, P.A., GDELT: Global data on events, location and tone, 1979-2012, *ISA Annual Convention*, 2013.x

[25] Gerner, D, Schrodt, P.,Abu-Jabr, R., and Yilmaz, O., Conflict and mediation event observations (CAMEO): A new event data framework for the analysis of foreign policy interactions. Working paper, 2012.

[26] L. Bottou, J. Peters, J. Quiñonero-Candela, D. X. Charles, D.M. Chickering, E. Portugaly, D. Ray, P. Simard, E. Snelson, "Counterfactual reasoning and learning systems: the example of computational advertising,", *Journal of Machine Learning Research*, vol. 14, no. 65, pp. 3207-3260, 2013.

[27] A. Gilotte, C. Calauzenes, T. Nedelec, A. Abraham, S. Dolle, "Offline A/B testing for recommender systems," in *Proc. of ACM International Conference of Web Search and Data Mining*, 2018.

[28] Choi, Y. and Wiebe, J., +/-EffectWordNet: Sense-level lexicon acquisition for opinion inference, *Proc. of the Conf. on Empirical Methods in Natural Language Processing (EMNLP)*, 2014.

[29] Pearl, J., Interpretation and Identification of Causal Mediation, *Psychol Methods*, vol. 19, no. 4, pp. 459-81, 2014.

[30] Granger, C. W. J., Investigating causal relations by econometric models and crossspectral methods, *Econometrica*, Vol 37, pp. 424–438, 1969.

[31] Friston, K.J., Bastos, A.M., Oswal, A., et al., Granger causality revisited, *Neuroimage*, Vol. 101, 2014.

[32] Jiang, L and Mai, F, Discovering bilateral and multilateral causal events in GDELT. Data challenge paper presented at the international conference on social computing, behavioral-cultural modeling, and prediction, Washington, DC, April 2–4, 2014

[5] Source code: https://github.com/lionelc/GDELT_Causal_Discovery