Causal Inference for Influence Propagation–Identifiability of the Independent Cascade Model

Shi Feng[1] Wei Chen[2]

[1]Institute for Interdisciplinary Information Sciences
Tsinghua University

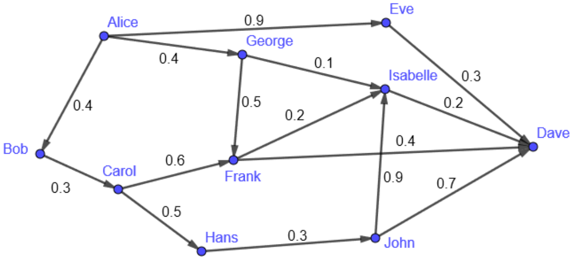
[2]Microsoft Research

CSoNet, November 2021

Overview

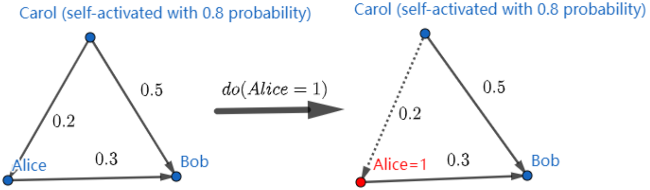
How to Model Diffusion in a Social Network?

- Traditionally, we use the **Independent Cascade model** (IC model).
- If it is a direct acyclic graph, it is exactly a Bayesian causal graph!
- After the propagation process, denote the activating status of node V_i by v_i.
- Suppose, $Pa(V_i) = \{V_{i_1}, V_{i_2}, \cdots, V_{i_k}\}$, we have $P(V_i = 1 | V_{i_1} = v_{i_1}, \cdots, V_{i_k} = v_{i_k}) = 1 - \prod_{j=1}^{k}(1 - p_{V_{i_j}, V_i} \cdot v_{i_j})$.



What is do effect (for modeling intervention)?

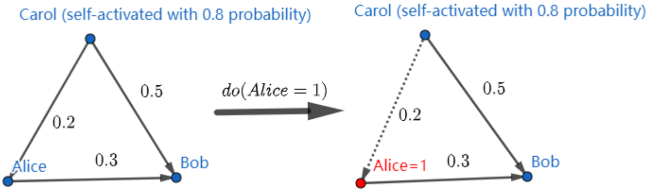
- Consider an IC model with three nodes, Salesman (Carol), Alice and Bob.
- Carol sold both Alice and Bob on the new operating system, Windows 11.
- If Alice bought it, she would recommend Bob to buy it.
- Node activation means purchase, non-activation means no purchase, and the activation probability is shown in the figure.



- $P(Bob = 1) = 1 - 0.8(1 - 0.2 \times 0.3)(1 - 0.5) = 0.53.$
- $P(Bob = 1|do(Alice = 1)) = 1 - 0.8(1 - 0.3)(1 - 0.5) = 0.65.$

What is do effect (for modeling intervention)?

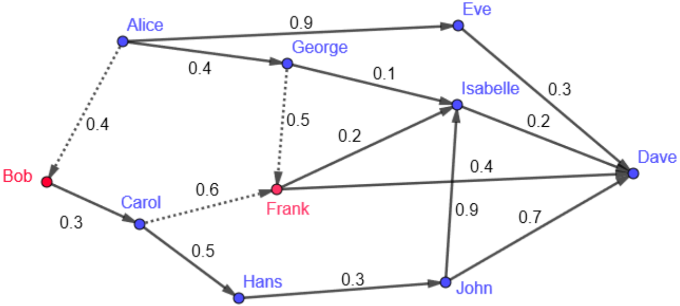
- $P(Bob = 1) = 1 - 0.8(1 - 0.2 \times 0.3)(1 - 0.5) = 0.53.$
- $P(Bob = 1|do(Alice = 1)) = 1 - 0.8(1 - 0.3)(1 - 0.5) = 0.65.$



- $do(Alice = 1)$ is an intervention that forcing Alice to use Windows 11, i.e. giving a free sample to her, which is different from $P(Bob = 1|Alice = 1)$.
- Actually, $P(Bob = 1|Alice = 1) = P(Bob = 1, Alice = 1)/P(Alice = 1) = 0.16(1 - 0.5 \times 0.3)/0.16 = 0.85.$

Seed Node Selection $\Longleftrightarrow$ Intervention

- Usually, we will choose a seed node set (Bob, Frank in our example).
- In-edge of Bob and Frank will be useless, Bob and Frank will be activated no matter how the propagation performs.
- Equivalent to the definition of *do*($\textit{Bob} = 1, \textit{Frank} = 1$).



Bayesian Causal Graph

- We can use Bayesian causal graph to model a social network!
- The propagating rule is equivalent to the Bayesian propagating rule if we merely observe the propagating results.
- $P(V_1 = v_1, \cdots, V_n = v_n) = \prod_{i=1}^{n} P(V_i = v_i | Pa(V_i) = pa(v_i))$.
- To be more specific, what we have is a form like this:

Probability \ Carol / Alice, Bob	Activated	Not Activated
A, A	0.1	0.2
A, N	0.1	0.05
N, A	0.3	0.15
N, N	0.05	0.05

How About IC Model that is not a DAG?

- The state of V_i in round t is $V_{i,t}$ and that $V_{i,t}$ has three values, 0, 1 and 2, for three states.
- State 0 means that the node is not activated.
- State 1 means that the node was activated at the last time point.
- State 2 means that the node is activated and has already tried to activate its child nodes.

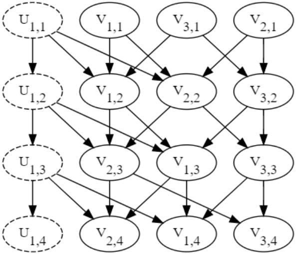


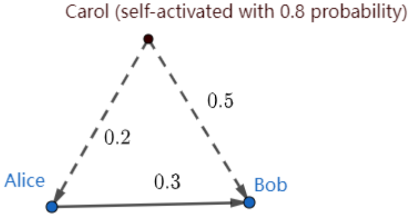
Figure: An example of transformation from IC model to Bayesian causal graph.

Identifiability Problem In Causal Graphs

- Now we have shown that how to transform an IC model to a causal graph.
- **Identifiability** generally says that if we can get the propagation result for infinitely times, we can completely restore the parameters in this graph.
- In Bayesian causal graph, **identifiability** means that if $P(V_1 = v_1, \cdots, V_n = v_n)$'s are known, we can solve $P(\mathbf{Y}|do(\mathbf{X} = \mathbf{x}))$ for node sets $\mathbf{X}, \mathbf{Y} \subset \mathbf{V}$.
- In IC model, **parameter identifiability** means that if $P(V_1 = v_1, \cdots, V_n = v_n)$'s are known ($2^n$ terms), we can solve all the activating probabilities p_{V_i,V_j} for $(V_i, V_j) \in \mathbf{E}$.
- With all the parameters, do effects can be naturally computed, so **parameter identifiability** is strictly stronger than **identifiability**!

With Hidden Variables

- If all the nodes are observable, all the do effects will be identifiable.
- If some variables are not observable? We still consider the Windows 11 selling example.
- We do not know Carol so we cannot know whether $Carol = 1$. Also, "Carol" can be a factor, such as common interests that cannot be revealed.
- So Carol is a hidden variable, the outcoming edges are denoted using dashed vectors.



- If we can only observe $P(Alice, Bob)$, can be get $P(Bob|do(Alice))$?

Identifiability Problem In Causal Graphs

- Fully solved for Semi-Markovian graphs!
- Pearl's do calculus algorithm is complete for Semi-Markovian Models[1].
- That is to say, after iterations of three rules in do calculus, if we can identify all the do effects, the causal graph is identifiable.

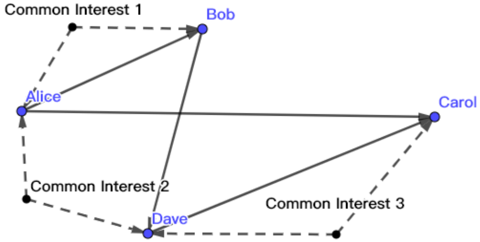
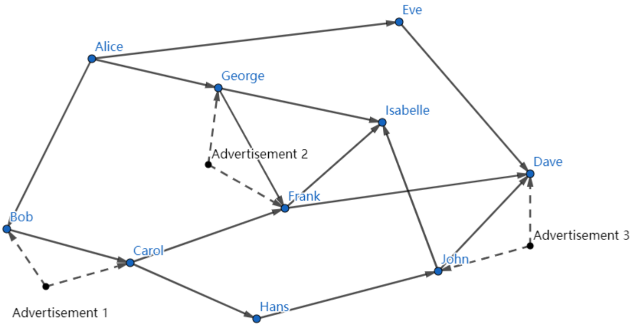


Figure: an example of semi-Markovian model, each hidden variable has at most two children and they should be observable

[1] Huang, Y., & Valtorta, M. (2012). Pearl's calculus of intervention is complete. arXiv preprint arXiv:1206.6831.

Identifiability Problem In IC Model

- Suppose $G = (\mathbf{U}, \mathbf{V}, \mathbf{E})$ where $\mathbf{U}, \mathbf{V}$ are the sets of hidden and observable variables, respectively.
- If we have $P(V_1, V_2, \cdots, V_n)$, can we solve all the parameters under any instance (taking any values for the parameters) of this graph G?
- For example, there are three hidden advertisements in our previous IC model.



Overview

Three Studied Typical IC Models

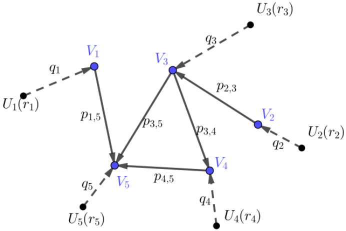


Figure: Markovian IC Model.

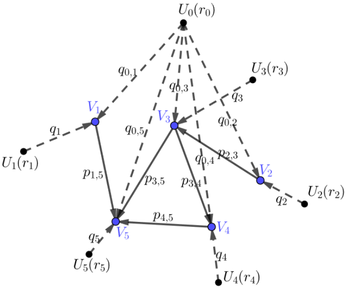


Figure: Markovian IC Model with a Global Effect.

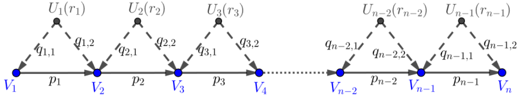
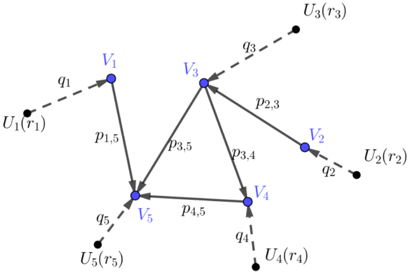


Figure: Semi-Markovian IC Model.

Markovian IC Models

Identifiability of the Markovian IC Model

For an arbitrary Markovian IC model $G = (U, V, E)$ with parameters $\boldsymbol{q} = (q_i)_{i \in [n]}$ and $\boldsymbol{p} = (p_{i,j})_{(V_i, V_j) \in E}$, all the q_i parameters are efficiently identifiable, and for every $i \in [n]$, if $q_i \neq 1$, then all $p_{j,i}$ parameters for $(V_j, V_i) \in E$ are efficiently identifiable.

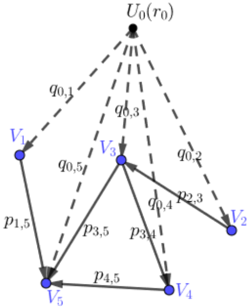


So almost all the Markovian IC models are identifiable!

IC Model with a Global Effect

Identifiability of the IC Model with a Global Hidden Variable

For an arbitrary IC model with a global hidden variable $G = (U, V, E)$ with parameters $\mathbf{q} = (q_i)_{i \in [n]}$, r and $\mathbf{p} = (p_{i,j})_{(V_i, V_j) \in E}$ such that $q_i \neq 1, p_{i,j} \neq 1$ and $r \neq 1$ for $\forall i, j \in [n]$, all the parameters in $\mathbf{p}$, r and $\mathbf{q}$ are identifiable.



So almost all the IC models with a global effect are identifiable!

Markovian IC Model with a Global Effect

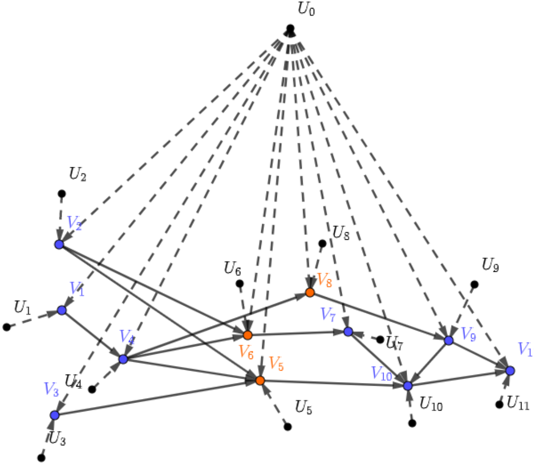
Identifiability of Markovian IC Model with a Global Hidden Variable

For an arbitrary Markovian IC Model with a Global Hidden Variable $G = (U, V, E)$ with parameters r_0, $\mathbf{q}_0 = (q_{0,i})_{i\in[n]}$, $\mathbf{q} = (q_i)_{i\in[n]}$ and $\mathbf{p} = (p_{i,j})_{(V_i,V_j)\in E}$, we suppose that all the parameters are not 1. If $\exists i, j, k \in [n], i < j < k$ such that each pair in V_i, V_j, V_k are disconnected and $q_{0,i}, q_{0,j}, q_{0,k} \neq 0$, then the parameters $q_{0,t}, q_t$ and $p_{t,l}, l > t > k$ are identifiable. Moreover, if V_i, V_j, V_k can be adjacently continuous in some topological order, i.e. $j = i + 1, k = i + 2$ without loss of generality, all the parameters are identifiable.

Comment: **unconnected** V_i, V_j, V_k **are the key!** If the tuple can be found, most parameters can be solved out; if the tuple contains continuous node in some topological order, all parameters can be solved out.

Markovian IC Model with a Global Effect

- If V_i, V_j, V_k can be adjacently continuous in some topological order, i.e. $j = i + 1, k = i + 2$ without loss of generality, all the parameters are identifiable. For example, V_5, V_6, V_8 in the figure. (idea: these three nodes can help us to speculate the state of global effect U_0.)



Unidentifiability of the Semi-Markovian IC Model

Unidentifiability of the Semi-Markovian IC Model

Suppose in a general graph G, we can find the following structure. There are three observable nodes V_1, V_2, V_3 such that $(V_1, V_2) \in E, (V_2, V_3) \in E$ and unobservable U_1, U_2 with $(U_1, V_1), (U_1, V_2), (U_2, V_2), (U_2, V_3) \in E$. Suppose each of U_1, U_2 only has two edges associated to it, the three nodes V_1, V_2, V_3 can be written adjacently in a topological order of nodes in $U \cup V$. Then we can deduce that the graph G is not parameter identifiable.

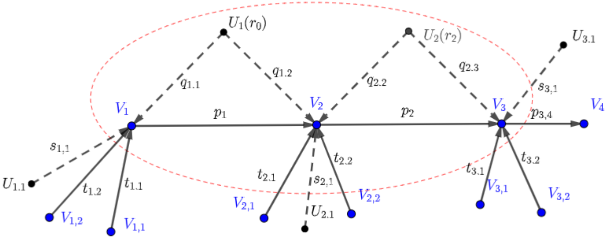
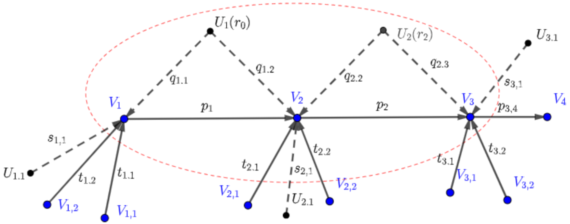


Figure: An example of such structure.

Unidentifiability of the Semi-Markovian IC Model

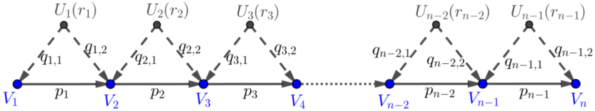
- The main idea is that construct two set of different parameters but they induce two same distribution on observable nodes.
- $r_1 = \frac{10r_2-7}{12r_2-10}, q_{1,1} = \frac{1}{4r_1}, q_{1,2} = \frac{6r_2-5}{8r_2-8}, q_{2,1} = \frac{1}{3-2r_2}, q_{2,2} = \frac{1}{4r_2}$ and other variables are fixed.
- This surprising construction is found by the method of undetermined coefficients.



So a large proportion of the semi-Markovian IC model is not fully identifiable!

Identifiability of the Semi-Markovian IC Model

- As we mentioned, the identifiability of semi-Markovian graphs is fully investigated in the context of causal graphs.
- What about the identifiability of the semi-Markovian IC model? Let us consider the simplest case - the chain.

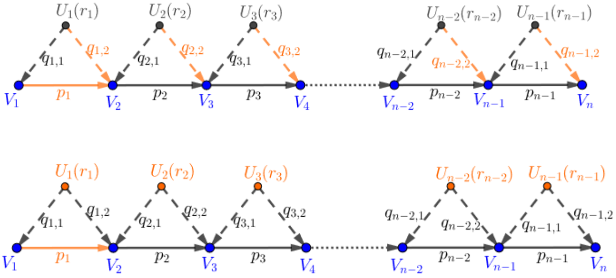


- However, according to our previous theorem, this contains $[\frac{n}{2}]$ of its required structure.
- So we prove that with n specific parameters known in advance, we can solve out all the other parameters.

Identifiability of the Semi-Markovian IC Model

Identifiability of Semi-Markovian IC Chain

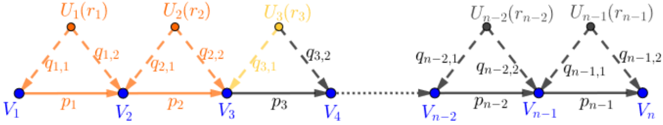
Suppose that we have a semi-Markovian IC chain model with the graph $G = (U, V, E)$ and the IC parameters $\boldsymbol{p} = (p_i)_{i \in [n-1]}$, $\boldsymbol{q}_1 = (q_{i,1})_{i \in [n-1]}$, $\boldsymbol{q}_2 = (q_{i,2})_{i \in [n-1]}$ and $\boldsymbol{r} = (r_i)_{i \in [n-1]}$, and suppose that all parameters are in the range $(0, 1)$. If the values of parameter p_1 is known, $\boldsymbol{q}_2$ or $\boldsymbol{r}$ is known, then the remaining parameters are efficiently identifiable.



Orange parameters are known in advance, then other parameters are identifiable.

Identifiability of the Semi-Markovian IC Model

- Proof idea: using mathematical induction, assume $p_1, p_2, \cdots, p_{t-2}$, $r_1, r_2, \cdots, r_{t-2}$, $q_{1,1}, q_{2,1}, \cdots, q_{t-2,1}$ and $q_{1,2}, q_{2,2}, \cdots, q_{t-2,2}, r_{t-1}q_{t-1,1}$ are already known.
- We prove that $q_{t-1,1}, r_{t-1}, p_{t-1}, q_{t-1,2}$ and $r_t q_{t,1}$ can be computed using distributions of observed variables.



- Solve the group of equations deduced by the parameter expression of $P(V_1 = 0, \cdots, V_{t-3} = 0, V_{t-2} = 1, V_{t-1} = 1, V_t = 0)$, $P(V_1 = 0, \cdots, V_{t-3} = 0, V_{t-2} = 0, V_{t-1} = 1, V_t = 0)$, $P(V_1 = 0, \cdots, V_{t-3} = 0, V_{t-1} = 0, V_t = 0)$. We have proved that other equations deduced by the distribution of first t nodes are all equivalent to these three.
- Actually, that is enough (but lots of technical issues not covered here).

Overview

Conclusions

Our Contributions:

- Proposed a method to convert the IC model into a Bayesian causal graph.
- Studied the problem of identifiability of IC models in detail, and give rich conditions for several types of common models to be identifiable as well as unidentifiable.
- Incorporation of observed confounding factors and causal inference techniques.

Future works

- Seed selection and influence maximization correspond to the intervention (or do effect) in causal inference.
- How to compute such intervention effect under the network with unobserved confounders and how to do influence maximization.
- Identifiability of the intervention effect, or whether given some intervention effect one can identify more of such effects.
- Identifiability in the general cyclic IC models ·

Q&A

Thank you for your attention!