

Distribution Grid Topology Reconstruction: An Information Theoretic Approach

Yizheng Liao, Yang Weng

Dept. of Civil and Environmental Engineering Dept. of Electrical Engineering Dept. of Civil and Environmental Engineering
Stanford University Stanford University Stanford University
Stanford, CA 94305 Stanford, CA 94305 Stanford, CA 94305

Email: {yzliao,yangweng}@stanford.edu

Meng Wu

Email: mengwu@stanford.edu

Ram Rajagopal

Email: ramr@stanford.edu

Abstract—Recently, a rapid penetration of distributed generation raises various issues. One of the key issues is frequent distribution grid re-configuration, which is hard to detect based on traditional approaches. Wrong topology information causes wrong control signal, making fast changing smart grid prone to go over stability boundaries and to collapse. To ensure system robustness, we propose a new data-driven re-configuration approach, thanks to recently progressively deployed larger sensor networks in distribution systems by utilities. Specifically, an Information Theory based algorithm, Chow-Liu algorithm, is used based on a proof for assumption and verification in distribution systems. Simulation results show highly accurate re-configuration estimation in IEEE standard distribution test systems.

Keywords—Power Distribution Network, Topology/Structure Learning, Information Theory, and Voltage Measurement

I. INTRODUCTION

Initiated by the U.S. government, the rapid-expanding smart grid aims at evolving into a sustainable modern grid. For this purpose, distributed energy resources (DER), such as photovoltaic and storage devices, are rapidly integrated into the distribution power grid for renewable generation. This is because generating power inside distribution power grid can not only create more sustainable energy sources, but can also create cheaper electricity and reduce losses due to shortened path between generation and end consumer, etc.

While producing new opportunities, the large-scale penetration of distributed generation is also posing new challenges. Different than transmission power grid where topology is with limited changes, a distribution grid can have regular topology changes due to ad-hoc connection of many plug-and-play components. Even worse, a distribution system operator usually lacks specific topology information, e.g., DG connection status, as many of the DERs do not belong to the utility.

Wrong topology estimate can cause critical issues such as wrong calculation in the dangerous reversions of power flow, incorrect description of fast dynamic variation of voltage profiles, and line work hazards, etc. Therefore, a highly active and accurate topology estimation process [1] is necessary in providing bases for distribution automation in the operation and control of smart distribution grid. This calls for an automatic, reliable, and low cost way to detect status changes.

Systematic topology error identification methods exist in the transmission grid based on a post-state estimation (SE) procedure [2]. Specifically, a topology error is detected if measurements associated with a branch or a bus are flagged as outliers by a SE-based residual test. In traditional transmission network, it is possible to use the topology identification process described above, based on the belief that no significant topology change appears in a short time with limited non-

conforming errors. However, such a belief will no longer hold in smart grid, where frequent topological changes appear, which may lead to large amount of topology errors, making the method above hard to identify all of them [3]–[5].

To resolve the problems, one can let utility company install monitoring devices on all topology changing components (e.g., switches) to achieve distribution automation. But this process will take a long time. One can also assume the availability of all possible switch connectivity map and search for the right combination [6]–[8]. Additionally, state estimation can be employed by assuming the availability of admittance matrix [9], [10]. Finally, power flow examination can be used with known admittance matrix [11].

Unfortunately, methods above require the knowledge of circuit breakers and/or admittance matrix, which may be unavailable in newly added or reconfigured partial distribution network. Even if there is such knowledge, it may be outdated or wrong due to human interaction without information updating, i.e. plug-and-play components in the distribution grids. This makes those methods unable to deal with large topology uncertainties in many smart grid scenarios.

Nowadays, the most reliable information usually come from the smart sensors at household level, thanks to recent advances in communications, sensing, computing and control, as well as the targeted investments toward deploying advanced meter infrastructures (AMIs) and synchrophasors, creating drivers and sources of data previously unavailable in the electric power industry [12], [13]. Therefore, in this paper, we restrict us to have only end-user data [14], [15], such as real & reactive powers and voltage magnitude measured by smart meter [16]–[19]. We also conduct analysis with voltage phase measured for the future when micro phasor measurement unit (μ PMU) [20], [21] has been widely installed. We aim at helping these devices identify the physical system in which they operate, discover their neighbor buses, and reconfigure the communication and control infrastructure in order to being able to perform assigned tasks.

Instead of assuming the knowledge of circuit breaker conditions or estimating system admittance matrix [3], we only use pairwise mutual information of voltage phasors. Historical data are used for mutual information calculation instead of using only a single data point (current method) [22]. We prove that voltage phasors data can be used to find the optimal topology connection via the Chow-Liu algorithm, enabling the identification of a large amount of plug-and-play devices based on their local sensor data against the ever-changing hard-to-predict uncertainties in smart grids.

Whereafter, the performance of the data-driven method is

verified by simulations on the standard IEEE 8- and 123-bus distribution test cases [23], [24]. Provided with enough historical data, the data-driven topology estimate outperforms the estimate from the traditional approach and approaches in recent paper [3]. As a highlight, error propagation in other approaches does not happen in our approach, due to the fact that our detection does not depend on known topology.

The rest of the paper is organized as follows: Section II introduces the modeling and the problem of data-driven topology identification. Section III uses a proof to justify the applicability of the Chow-Liu algorithm for distribution system topology re-configuration. A detailed algorithm is illustrated as well. Section IV evaluates the performance of the new method and Section V concludes the paper.

II. SYSTEM MODELS

To better formulate the topology reconstruction problem, we need to describe the distribution network topology and data. A distribution network is characterized by *buses*, $\mathcal{V} = \{1, 2, \dots, N\}$, and by *branches* (power lines), $\mathcal{E} = \{(i, j), i, j \in \mathcal{V}\}$. Measurement data at bus i and time t can be represented as follows: active power $p_i(t) \in \mathbb{R}$, reactive power $q_i(t) \in \mathbb{R}$, and voltage $v_i(t) = |v_i(t)|e^{j\theta_i(t)} \in \mathbb{C}$, where $|v_i(t)| \in \mathbb{R}$ denotes the voltage magnitude in per unit and $\theta_i(t) \in \mathbb{R}$ denotes the voltage phase angle in degree. These measurements are in the steady state and all voltages and currents are sinusoidal signals at the same frequency.

The problem of distribution grid topology reconstruction is defined as follows:

- Problem: data-driven topology reconstruction based on voltages
- Given: a sequence of historical measurements $v_i(t), i \in \mathcal{V}, t = 1, \dots, T$ and a partially known grid topology, as shown in Fig. 1
- Find: the local grid topology \mathcal{E} in the dashed box in Fig. 1

III. INFORMATION THEORY-BASED TOPOLOGY ESTIMATION

Recent deployed smart meters provide a large amount of highly accurate time-series data. We want to utilize these data to reconstruct grid topology. One way to represent the historical data is using a probability distribution. For example, the joint distribution of voltage measurements is

$$P(\mathbf{V}) = P(V_2, V_3, \dots, V_N) \\ = P(V_2)P(V_3|V_2) \dots P(V_N|V_2, \dots, V_{N-1}), \quad (1)$$

where $V_i \in \mathbb{C}$ is a continuous complex random variable that represents the voltage measurement $v_i(t)$. Bus 1 is omitted from this joint distribution because it is the slack bus with fixed voltage measurement over time, i.e. $1 \angle 0$. A distribution system network usually has a radial structure. Also, the correlation between interconnected neighboring buses are higher than that of non-neighbor buses. Therefore, a reasonable approximation of (1) is to assume that a nodal measurement depends on its neighbors' measurements. Subsequently, (1) is simplified as

$$P(\mathbf{V}) \approx P_t(\mathbf{V}) = \prod_{i=2}^N P(V_i|V_{r(i)}), \quad (2)$$

where $r(i)$ is the parent bus¹ that connects with bus i . $P(V_2|V_{r(2)})$ is defined to be equal to $P(V_2)$. Such simplification is called *product approximation*. After this definition,

¹In a tree graph, the parent bus is defined as the bus that is closer to the feeder.

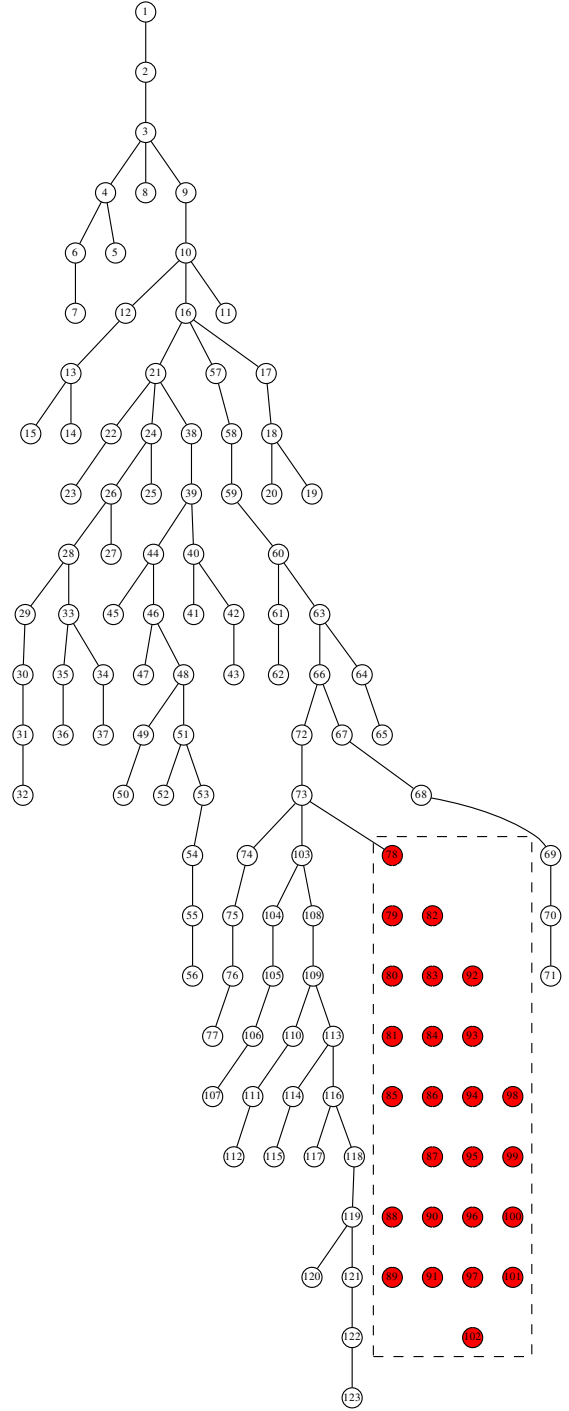


Fig. 1. 123-bus system: the framework of data-driven state estimation.

finding the grid topology is now equivalent to finding the conditional distribution $P_t(\mathbf{V})$ that best approximate the joint distribution $P(\mathbf{V})$.

The Chow-Liu algorithm has been shown as an optimal algorithm that finds the best product approximation of $P(\mathbf{V})$ [25]. It uses the mutual information of all possible bus pairs within the network and finds maximum weight spanning tree that maximizes the overall mutual information. Mutual information, an information theory metric that measures the mutual dependence of two random variables, has been widely used in many fields, such as communication [26], natural

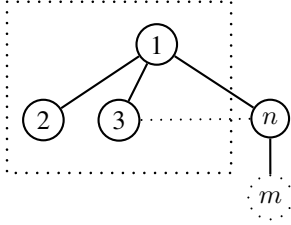


Fig. 2. $(n+1)$ -bus system: depth is 3.

language processing [27], and biomedical engineering [28]. Mutual information measures how similar the joint distribution of two random variables, $p(V_i, V_j)$, is to the products of the individual distributions, $p(V_i)p(V_j)$. For continuous random variables, mutual information is defined as [29]

$$I(V_i, V_j) = \int_{V_i} \int_{V_j} p(v_i, v_j) \ln \left(\frac{p(v_i, v_j)}{p(v_i)p(v_j)} \right) dv_i dv_j. \quad (3)$$

Alternatively, mutual information can be computed using entropy as

$$I(V_i, V_j) = H(V_i) + H(V_j) - H(V_i, V_j), \quad (4)$$

where the entropy $H(V_i)$ is defined as

$$H(V_i) = - \int_{V_i} p(v_i) \ln(p(v_i)) dv_i,$$

and $H(V_i, V_j)$ denotes entropy of the joint distribution [29]. In Theorem 1, we prove that the Chow-Liu algorithm is suitable for the distribution power grid.

Theorem 1. *In a distribution power grid with tree structure, the Chow-Liu algorithm finds the optimal product approximation of $P(\mathbf{V})$ and its associated topology connection.*

Proof: Here we use a 3-bus system (the dashed box in Fig. 2) to prove the conditional independence of voltages, i.e. $V_2|V_1 \perp V_3|V_1$, which serves as the basis of applying the Chow-Liu algorithm.

Let the current injection $I_i \in \mathbb{C}$ be a random variable for $i = 1, 2, 3$ and be independent with others. Also, let $y_{ij} \in \mathbb{C}$ denote the line admittance between bus i and bus j . $y_{ij} = 0$ when no single branch exists between two buses. Therefore, the relationship between voltages and currents is:

$$\begin{bmatrix} y_{12} + y_{13} & -y_{12} & -y_{13} \\ -y_{12} & y_{12} & 0 \\ -y_{13} & 0 & y_{13} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix}.$$

When $V_1 = v_1$, rewriting the relationship above leads to

$$\begin{bmatrix} -y_{12} & -y_{13} \\ y_{12} & 0 \\ 0 & y_{13} \end{bmatrix} \begin{bmatrix} V_2 \\ V_3 \end{bmatrix} = \begin{bmatrix} I_1 - v_1(y_{12} + y_{13}) \\ I_2 + v_1 y_{12} \\ I_3 + v_1 y_{13} \end{bmatrix}.$$

As the first row is redundant, the linear system turns to

$$V_2 = \frac{I_2}{y_{12}} + v_1, \quad V_3 = \frac{I_3}{y_{13}} + v_1. \quad (5)$$

Since I_2 and I_3 are independent, V_2 and V_3 are independent conditioning on V_1 , i.e. $V_2|V_1 \perp V_3|V_1$.

Next, we extend the conditional independence to a more general network. For the grid shown in Fig. 2 (excluding Bus m), there are n buses. $n-1$ buses have a common neighbor, namely Bus 1. Using similar approach as the 3-bus system example, given $V_1 = v_1$, we can find the following equation for each bus i ,

$$V_i = \frac{I_i + v_1 y_{1i}}{y_{ii}}, \quad i = 2, \dots, n. \quad (6)$$

This equation has a consistent format as (5) since $y_{ii} = y_{1i}$ in current setup. With the same assumption that each current

injection I_i is independent with others, we can conclude that $V_i|V_1$ is independent with $V_j|V_1$ for $i, j = 2, \dots, n$ and $i \neq j$.

When Bus m is attached to Bus n , the new relationship between currents and voltages is

$$\begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \\ I_m \end{bmatrix} = \begin{bmatrix} y_{11} & -y_{12} & \dots & -y_{1n} & 0 \\ -y_{12} & y_{22} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -y_{1n} & 0 & \dots & y_{nn} & -y_{nm} \\ 0 & \dots & \dots & -y_{nm} & y_{mm} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \\ V_m \end{bmatrix}.$$

Given $V_1 = v_1$, we have

$$\begin{bmatrix} I_2 + v_1 y_{12} \\ \vdots \\ I_n + v_1 y_{1n} \\ I_m \end{bmatrix} = \begin{bmatrix} y_{22} & 0 & \dots & 0 & 0 \\ 0 & y_{33} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & y_{nn} & -y_{nm} \\ 0 & 0 & \dots & -y_{nm} & y_{mm} \end{bmatrix} \begin{bmatrix} V_2 \\ V_3 \\ \vdots \\ V_n \\ V_m \end{bmatrix}.$$

For Bus 2 to Bus $n-1$, the proof in the first part, i.e. (6), still holds. The only exception is determining if the tree dependence still holds for Bus n , which is now connected with Bus m . To explore the conditional independence, between Bus n and Bus $i \in \{2, \dots, n-1\}$, we need to use the following relationship extracted from the formula above.

$$I_n + v_1 y_{1n} = y_{nn} V_n - y_{nm} V_m, \quad (7)$$

$$I_m = -y_{nm} V_n + y_{mm} V_m. \quad (8)$$

Since Bus m only connects with Bus n , $y_{mm} = y_{nn}$. Further, $y_{nn} = y_{nm} + y_{1n}$ by the definition of admittance matrix. Hence, by combining (7) and (8), we have

$$\begin{aligned} I_n + v_1 y_{1n} + I_m &= y_{nn} V_n - y_{nm} V_m \\ &\quad - y_{nm} V_n + y_{mm} V_m \\ &= y_{1n} V_n, \\ V_n &= \frac{I_n + I_m + v_1 y_{1n}}{y_{1n}}. \end{aligned} \quad (9)$$

Since the current injection is independent at each bus, $I_i \perp I_j$ for $2 \leq i \leq n-1$ and $j = n, m$. Then, $I_n + I_m$ is independent with I_i as well. Therefore, $V_n|V_1$ is independent with $V_i|V_1$ due to (6) and (9).

In conclusion, this proof can be easily extended to the case where each bus has one bus connected, other than Bus 1. If each bus has over one bus connected, we can aggregate these buses into a single bus and use the proof above to show the conditional independence of voltage phasors.

With the conditional independence of buses, via mutual information, the Chow-Liu algorithm finds the optimal approximation to $P(\mathbf{V})$, $P_t(\mathbf{V})$. ■

To illustrate the steps of the Chow-Liu algorithm, we summarize them in a flow chart in Fig. 3 as well as an algorithm table at the end of this section.

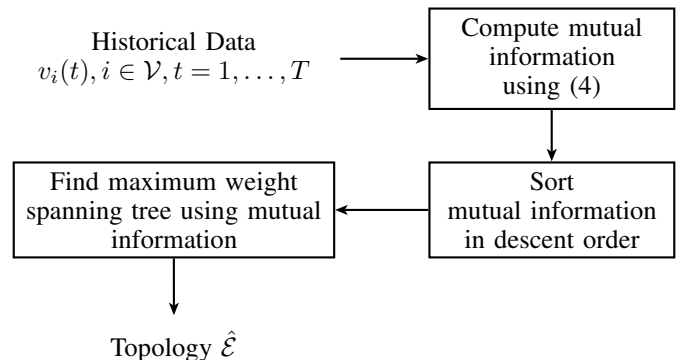


Fig. 3. Flow chart of the proposed approach.

One key step in the flow chart above is to compare the mutual information. We use the following lemma to illustrate why such concept is important to find the correct topology.

Lemma 1. *In a distribution network with tree structure and conditional independence assumption in Theorem 1, $I(V_j, V_i) \geq I(V_j, V_k)$ given $j, k \in r(i)$, $k \notin r(j)$ and $j \notin r(k)$.*

Proof:

$$\begin{aligned} I(V_i, V_j, V_k) &= I(V_i, V_j) - I(V_i, V_j|V_k) \\ &= I(V_j, V_k) - I(V_j, V_k|V_i). \end{aligned}$$

Since $V_j|V_i$ is independent with $V_k|V_i$, the conditional mutual information $I(V_j, V_k|V_i)$ is zero. Then we have

$$I(V_i, V_j) = I(V_j, V_k) + I(V_i, V_j|V_k).$$

Due to the fact that mutual information is always non-negative, we have

$$I(V_j, V_i) \geq I(V_j, V_k). \quad \blacksquare$$

According to Lemma 1, buses have much higher mutual information with their neighbors. In next section, we will use the numerical example to demonstrate this lemma. By using the mutual information as the weight and Theorem 1, a maximum weight spanning tree will find a highly accurate topology of a distribution grid, making the Chow-Liu algorithm suitable for the smart grid topology identification.

Algorithm 1 Distribution Grid Topology Reconstruction

Require: $v_i(t)$ for $i = 2, \dots, N$, $t = 1, \dots, T$

```

1: for  $i, j = 2, \dots, N$  do
2:   Compute mutual information  $I(V_i, V_j)$  using (4) based
   on  $v_i(t)$ .
3: end for
4: Sort all possible bus pair  $(i, j)$  into nonincreasing order
   by  $I(V_i, V_j)$ . Let  $\hat{\mathcal{E}}$  denote the sorted set.
5: Let  $\hat{\mathcal{E}}$  be the set of nodal pair comprising the maximum
   weight spanning tree. Set  $\hat{\mathcal{E}} = \emptyset$ .
6: for  $(i, j) \in \hat{\mathcal{E}}$  do
7:   if cycle is detected in  $\hat{\mathcal{E}} \cup (i, j)$  then
8:     Continue
9:   else
10:     $\hat{\mathcal{E}} \leftarrow \hat{\mathcal{E}} \cup (i, j)$ 
11:   end if
12:   if  $|\hat{\mathcal{E}}| == N - 2$  then
13:     break
14:   end if
15:   return  $\hat{\mathcal{E}}$ 
16: end for
```

Step 6-16 build a maximum spanning tree using pairwise mutual information as the weight. This algorithm is modified from the well-known Kruskal's minimum weight spanning tree algorithm [30], [31], which has a running time of $O((N-2)\log(N-1))$ for a radial distribution network with N buses. Therefore, the proposed algorithm can efficiently reconstruct the topology with low computational complexity. As our detection method does not depend on known topology, error propagation does not happen in our approach.

IV. SIMULATION AND RESULTS

The simulations are implemented on the IEEE PES distribution networks for IEEE 8-bus and 123-bus networks [32]. To better visualize the connection of distribution grid topology

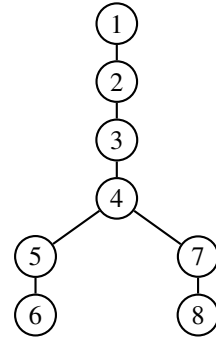


Fig. 4. 8-bus system.

and the graphical tree, we transform the original power grid maps to tree graphs in Fig. 1 and Fig. 4. In these figures, nodes represent buses and edges represent branches. In each network, Bus 1 is selected as the slack bus. The historical data have been preprocessed by the MATLAB Power System Simulation Package (MATPOWER) [23], [24]. To simulate the power system behavior in a more practical pattern, the load profile from PJM [33] is adopted as the real power profile in the subsequent simulation. The load data used are between 2008 and 2014 with a consistent data format. For the reactive power (q_i) at Bus i , we simulate it according to an independent and identical uniform distribution, i.e.

$$q_i(t) \sim \text{Unif}(0.5\mu_i^q, 1.5\mu_i^q) \quad t = 1, \dots, T,$$

where the mean μ_i^q is given in the IEEE PES distribution network. To obtain other measurements, i.e. $|v_i(t)|$ and $\theta_i(t)$, we run a power flow to generate the states of the power system. To obtain time-series data, we run the power flow to generate hourly data over a year. $T = 8736$ measurements are obtained at each bus. Finally, we only use the voltage measurement to run the Chow-Liu algorithm.

To simplify the analysis, we model V_i at Bus i as a two-dimensional real Gaussian random vector, instead of a complex random variable,

$$\begin{bmatrix} \mathcal{R}(V_i) \\ \mathcal{I}(V_i) \end{bmatrix} \sim \mathcal{N}(\mu_i, \Sigma_i),$$

where $\mathcal{R}(V_i)$ and $\mathcal{I}(V_i)$ denote the real and imaginary parts of V_i , $\mu_i \in \mathbb{R}^2$ denotes the mean vector and $\Sigma_i \in \mathbb{R}^{2 \times 2}$ denotes the covariance matrix. For a k -dimensional Gaussian random vector $\mathbf{X} \sim \mathcal{N}(\mu, \Sigma)$, the entropy is defined as

$$H(\mathbf{X}) = \frac{k}{2}(1 + \ln(2\pi)) + \frac{1}{2} \ln |\Sigma|.$$

In the simulation, the mean vector and the covariance matrix at each bus are learned from historical observations.

A. Mutual information for topology detection

Fig. 5 shows the heat map of the mutual information matrix for the IEEE 123-bus system. Since the mutual information matrix is symmetric ($I(X, Y) = I(Y, X)$), we only mark branches in the lower triangular part. Circle represents the true connection. Cross represents the connection between the row and column indexes. If a circle is superposed by a cross, a correct topology identification is claimed. From Fig. 5, we observe that

- the first row and the first column have zero mutual information because they are associated with the slack bus, which has fixed voltage phasor, i.e. $1 \angle 0$.

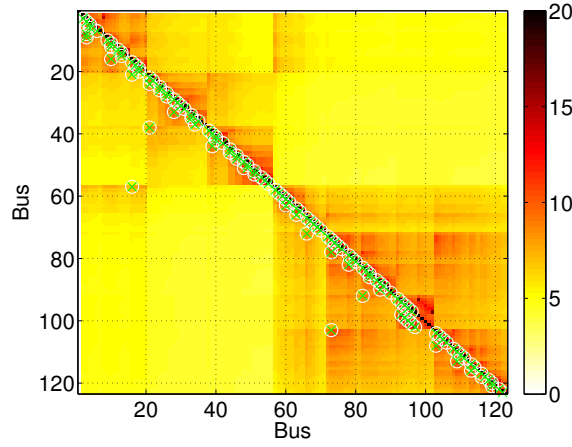


Fig. 5. Heat map of the mutual information matrix. The white circle indicates the branches in the 123-bus network. The green cross indicates the detected branches.

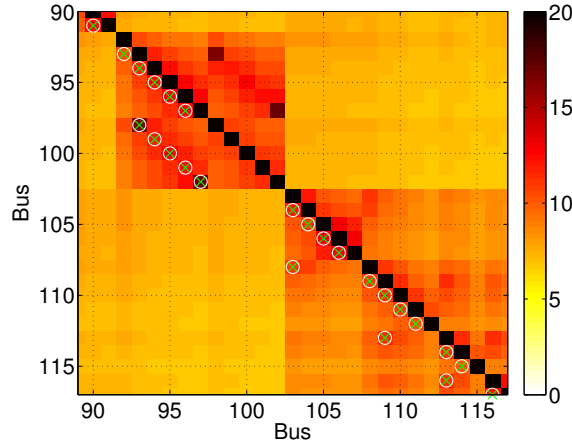


Fig. 6. Heat map of the mutual information matrix (Zoomed in). The white circle indicates the branches in the 123-bus network. The green cross indicates the detected branches.

- the diagonal element has the largest mutual information in each row because it is self-information [29].
- the coordinate associated with the true branch has the large mutual information in each row (excluding the diagonal element), which is zoomed-in inside Fig. 6. **This fact illustrates that using the pairwise mutual information as the weight is consistent with the physical behaviors.**

Fig. 7 displays pairwise mutual information of two specific buses, i.e. Bus 26 and Bus 109. We can clearly see that the mutual information of the nodal pairs in \mathcal{E} is very large compared with others.

B. Successful Detection Rate

To summarize performances in various simulation cases, we define the successful detection rate (SDR) as

$$\text{SDR} = \frac{\sum_{i,j \in \mathcal{E}} \mathbb{I}((i,j) \in \hat{\mathcal{E}})}{|\mathcal{E}|} \times 100\%,$$

where $\hat{\mathcal{E}}$ denotes the estimated set of branches and $|\mathcal{E}|$ denotes the size of the set \mathcal{E} . Table I summarizes SDR on 8-bus and 123-bus systems. As shown in the first column, the proposed algorithm can recover the entire 8-bus system without error. When we try to reconstruct the 123-bus system, SDR is still

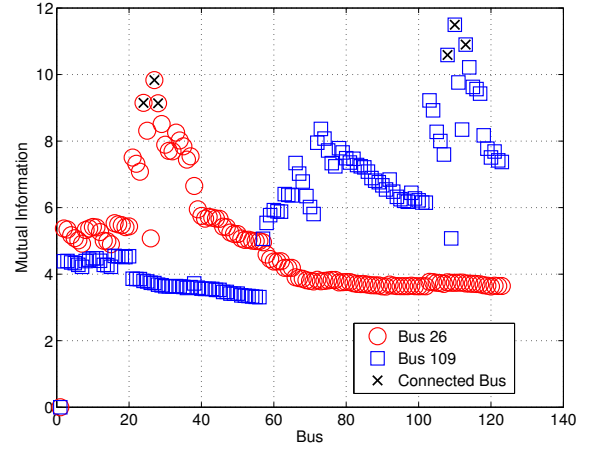


Fig. 7. Pairwise mutual information

TABLE I. SUCCESSFUL DETECTION RATE

	8-Bus Network	123-Bus Network
Historical P	100%	98.4%
Simulated P	100%	100%

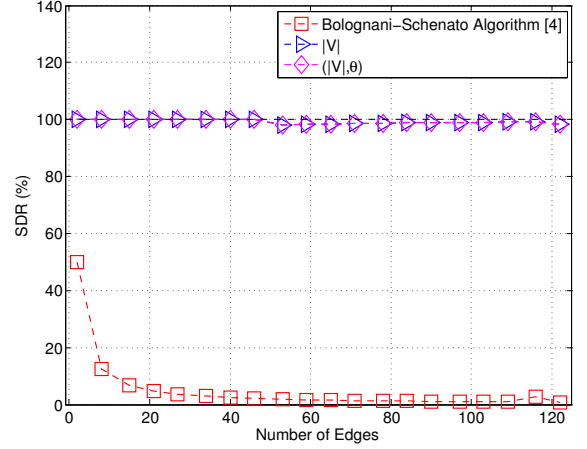


Fig. 8. SDR comparison of the proposed algorithm and the Bolognani-Schenato algorithm [3].

high, which is either 100% or 98.4%. Recovering a subnetwork is very practices in distribution networks, where only a subarea has unknown topology. We applied the algorithm to reconstruct a subtree of the large-scale system, i.e. 57 buses out of 123 buses, we can always reconstruct the topology correctly.

In Fig. 8, we compare the proposed algorithm with the algorithm in [3]. The x -coordinate represents the number of edges that are needed to be identified. The y -coordinate represents the successful detection rate. As the number of unknown edges increases, our approach consistently has a successful rate of 100%, while the other method's detection ability decreases. Additionally, we demonstrate that using voltage magnitude measurements can achieve similar results as using voltage phasor.

C. Sensitivity to the training data length

To explore how sensitive the propose algorithm is to the number of samples, we run Monte Carlo simulation by using 10 to 300 days data. The results are shown in Fig. 9. We observe that when more than 50 days' observations are available, our algorithm can stably reconstruct the topology

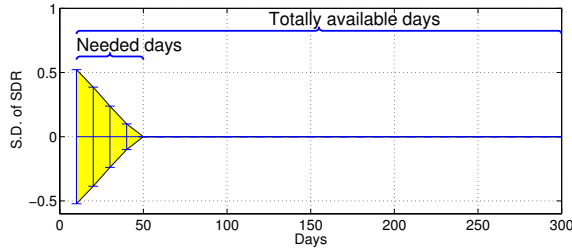


Fig. 9. Standard deviation of SDR against total number of observation.

with historical data. For the simulated power profile, our results show that only 20 days' data are needed. These results reflect that our algorithm can provide robust reconstruction.

V. CONCLUSION

In this paper, we propose a data-driven algorithm that reconstructs the topology of distribution networks. Unlike existing approaches, which require the knowledge about branches, our algorithm utilizes the newly available smart metering data only. We formulate the topology reconstruction problem as a joint distribution (voltage phasors) approximation problem. We prove that the optimal approximation can be achieved by applying the Chow-Liu algorithm via a comparison of pairwise mutual information. We verify the proposed algorithm on both IEEE 8- and 123-bus systems and results are highly accurate. With real load profile, we achieve zero reconstruction error on both 8-bus system and 123-bus system with partially known topology. Finally, our algorithm can perfectly reconstruct the partial topology with a limited amount of data.

ACKNOWLEDGMENT

The second author would like to thank Tie Liu from Texas A&M University for discussions on the Chow-Liu algorithm.

REFERENCES

- [1] A. Wood and B. Wollenberg, "Power generation, operation, and control, 2nd ed." Wiley, 1996.
- [2] R. Lugtu, D. Hackett, K. Liu, and D. Might, "Power system state estimation: Detection of topological errors," *IEEE Transactions on Power Apparatus and Systems*, 1980.
- [3] S. Bolognani, N. Bof, D. Michelotti, R. Muraro, and L. Schenato, "Identification of power distribution network topology via voltage correlation analysis," in *2013 IEEE 52nd Annual Conference on Decision and Control*. IEEE, 2013, pp. 1659–1664.
- [4] S. Xu, R. C. de Lamare, and H. V. Poor, "Dynamic topology adaptation for distributed estimation in smart grids," in *2013 IEEE 5th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*. IEEE, 2013, pp. 420–423.
- [5] J. Huang, V. Gupta, and Y.-F. Huang, "Electric grid state estimators for distribution systems with microgrids," in *2012 46th Annual Conference on Information Sciences and Systems*. IEEE, 2012, pp. 1–6.
- [6] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning in power distribution networks," *arXiv preprint arXiv:1501.04131*, 2015.
- [7] G. Cavraro, R. Arghandeh, and A. von Meier, "Distribution network topology detection with time series measurement data analysis," *arXiv preprint arXiv:1504.05926*, 2015.
- [8] Y. Sharon, A. M. Annaswamy, A. L. Motto, and A. Chakraborty, "Topology identification in distribution network with limited measurements," in *2012 IEEE Power and Energy Society General Meeting: Innovative Smart Grid Technologies*. IEEE, 2012, pp. 1–6.
- [9] G. N. Korres and N. M. Manousakis, "A state estimation algorithm for monitoring topology changes in distribution systems," in *2012 IEEE Power and Energy Society General Meeting*. IEEE, 2012, pp. 1–8.
- [10] M. Baran, J. Jung, and T. McDermott, "Topology error identification using branch current state estimation for distribution systems," in *2009 IEEE Transmission & Distribution Conference & Exposition: Asia and Pacific*. IEEE, 2009, pp. 1–4.
- [11] R. Arghandeh, M. Gahr, A. von Meier, G. Cavraro, M. Ruh, and G. Andersson, "Topology detection in microgrids with micro-synchrophasors," *arXiv preprint arXiv:1502.06938*, 2015.
- [12] Y. Weng, R. Negi, and M. Ilic, "Historical data-driven state estimation for electric power systems," *IEEE SmartGridComm Symposium*, Oct. 2013.
- [13] —, "Graphical model for state estimation in electric power systems," *IEEE SmartGridComm Symposium*, Oct. 2013.
- [14] Y. Weng and R. Rajagopal, "Probabilistic baseline estimation via gaussian process," *IEEE Power and Energy Society General Meeting*, July 2015.
- [15] J. Yu, Y. Weng, and R. Rajagopal, "Probabilistic power systems intervention estimation," *IEEE SmartGridComm Symposium*, 2015, accepted.
- [16] X. Han, S. You, F. Thordarson, D. V. Tackie, S. M. Ostberg, O. M. Pedersen, H. W. Bindner, and N. C. Nordentoft, "Real-time measurements and their effects on state estimation of distribution power system," in *4th IEEE/PES Innovative Smart Grid Technologies Europe*. IEEE, 2013, pp. 1–5.
- [17] A. Valdborg, C. Torchia, and M. Dwyer, "Application of southern california edison company (u 338-e) for approval of its distribution resources plan," Southern California Edison Company, July 2015.
- [18] P. P. System. Smart meter. [Online]. Available: <http://pacepowersystems.com/smart-meters/#.Va2IVRNviko>
- [19] M. Weiss, A. Helfenstein, F. Mattern, and T. Staake, "Leveraging smart meter data to recognize home appliances," in *IEEE International Conference on Pervasive Computing and Communications*. IEEE, 2012, pp. 190–197.
- [20] A. von Meier, D. Culler, A. McEachern, and R. Arghandeh, "Micro-synchrophasors for distribution systems," in *IEEE PES Innovative Smart Grid Technologies Conference*, Feb 2014, pp. 1–5.
- [21] G. Cavraro, R. Arghandeh, A. von Meier, and K. Poola, "Data-driven approach for distribution network topology detection," *arXiv preprint arXiv:1504.00724*, 2015.
- [22] Y. Weng, C. Faloutsos, and M. Ilic, "Data-driven topology estimation," *IEEE SmartGridComm Symposium*, Nov. 2014.
- [23] R. D. Z. C. E. Murillo-Sanchez and R. J. Thomas, "Matpower's extensible optimal power flow architecture," *IEEE Power and Energy Society General Meeting*, 2009.
- [24] R. D. Zimmerman and C. E. Murillo-Sanchez, "Matpower, a matlab power system simulation package," <http://www.pserc.cornell.edu/matpower/manual.pdf>, 2010.
- [25] C. Chow and C. Liu, "Approximating discrete probability distributions with dependence trees," *IEEE Transactions on Information Theory*, vol. 14, no. 3, pp. 462–467, 1968.
- [26] A. J. Goldsmith and P. P. Varaiya, "Capacity, mutual information, and coding for finite-state markov channels," *IEEE Transactions on Information Theory*, vol. 42, no. 3, pp. 868–886, 1996.
- [27] K. W. Church and P. Hanks, "Word association norms, mutual information, and lexicography," *Computational linguistics*, vol. 16, no. 1, pp. 22–29, 1990.
- [28] R. Moddemeijer, "On estimation of entropy and mutual information of continuous distributions," *Signal Processing*, vol. 16, no. 3, pp. 233–248, 1989.
- [29] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
- [30] J. B. Kruskal, "On the shortest spanning subtree of a graph and the traveling salesman problem," *Proceedings of the American Mathematical society*, vol. 7, no. 1, pp. 48–50, 1956.
- [31] T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein *et al.*, *Introduction to algorithms*. MIT press Cambridge, 2001, vol. 2.
- [32] W. H. Kersting, "Radial distribution test feeders," in *2001 IEEE Power Engineering Society Winter Meeting*, vol. 2. IEEE, 2001, pp. 908–912.
- [33] PJM. Pjm metered load data. [Online]. Available: <http://www.pjm.com/markets-and-operations/ops-analysis/historical-load-data.aspx>