Fault Localization in Power Distribution Networks using Operative Diagnosis

Girish Keshav Palshikar
Tata Research Development and Design Centre (TRDDC),
54B, Hadapsar Industrial Estate,
Pune 411013, India.
GK.Palshikar@tcs.com

ABSTRACT
Maintenance of power distribution networks is a critical task, which becomes complicated when parts of the network are not equipped with fault detection sensors. In this paper, we apply the operative diagnosis framework [9], [2], [11] to the problem of fault localization. In operative diagnosis, a system is prepared which represents the components and the flow of fault among components. Some of the components are equipped with alarms, which ring when a fault reaches that component. Given a system graph and a subset of ringing alarms, an operative diagnosis algorithm performs a root-cause analysis and identifies a collection of diagnoses (a diagnosis is a set of faulty components), each of which explains ringing and non-ringing alarms. We represent a power distribution network as a system graph, and customer complaints (of power outage) as ringing alarms and use operative diagnosis algorithms to locate faulty components. The operative diagnosis techniques can also be applied to systems other than power distribution networks.

23. INTRODUCTION
Electrical power distribution networks are a vital resource of a nation. Consequently, their maintenance is a critical task, which helps in ensuring the quality and reliability of the power distribution. Several factors complicate the task of maintenance of electrical power distribution networks. Typically, the size of the power distribution networks is very large and the circuit structure (i.e., interconnections among components) can be complex. Typically, the number of consumers (even within a region) is in millions, the number of equipment such as transformers and poles is in hundreds of thousands, the wire length is in tens of thousands of miles. Moreover the yearly number of distribution faults (major or minor) that occur is also often very large, typically in tens of thousands.

The power distribution networks suffer many different types of faults, both due to usual reasons such as wear and tear of the components and due to catastrophic conditions as in storms or lightning. Maintenance is then very critical in ensuring smooth operations of the network. Effective fault diagnosis in power distribution networks can help reduce the outage time and reduce the maintenance costs. Despite the complexity in the network size, the costs for maintaining the power distribution networks need to be kept as low as possible. Automation is one possible answer to achieve lower maintenance costs. There is a clear trend towards equipping the electrical power distribution networks with SCADA-enabled failure detection equipment, so that identifying the faulty components is quick, easy and accurate. However, such automation is expensive, particularly when carried to the lower levels of the distribution network (11kV or less), primarily due to the very large number of equipment that may need to be monitored. These problems (network size and lack of automation) are particularly acute for developing countries like India, where telemetry data is not available for low-voltage power distribution networks.

In this paper, we examine the problem of fault localization in power distribution networks, in the absence of detailed telemetry data. We then solve this problem in a well-known graph-theoretical framework for root-cause fault analysis called operative diagnosis. In operative diagnosis, a system is prepared which represents the components and the flow of fault among components. Some of the components are equipped with alarms, which ring when a fault reaches that component. Given a system graph and a subset of ringing alarms, operative diagnosis algorithms perform a root-cause analysis and identifies a collection of diagnoses (a diagnosis is a set of faulty components), each of which explains ringing and non-ringing alarms. We represent a power distribution network as a system graph, and customer complaints (of power outage) as ringing alarms and use operative diagnosis algorithms to locate faulty components. We assume that the customer complaints necessarily indicate presence of permanent faults (they are not about transient faults). The proposed approach helps in reducing the automation costs without compromising on the fault localization accuracy. We illustrate the approach on a small real-life network.

In section 2, we summarize the maintenance process for power distribution networks and motivate the problem. Section 3 contains an overview of the operative diagnosis framework. In section 4, we apply the operative diagnosis framework to solve the fault localization problem. We also briefly describe our prototype fault localization system based on this approach. Section 5 contains some related work. Section 6 contains our conclusions and states some future work.

24. MAINTENANCE PROCESS
The maintenance process for electrical power distribution networks consists of various activities. The first set of activities are related to the process of what one may call active maintenance, which is needed when a fault occurs when the system is in active use by the users and needs to be corrected as
soon as possible. The active maintenance process consists of activities such as the following:

- **fault detection**: establishing that a fault – such as power outage – has occurred. Most typical symptom of a fault in an electrical power distribution network is of course power outage to some of its consumers.

- **fault localization**: identifying the primary faulty components (e.g., their identity and physical location) whose failure has lead to the occurrence of the fault

- **fault diagnosis**: identifying the reasons for the failure of the components. This is typically a very technical activity that has to be done at the site of (each of) the faulty component and may involve detailed knowledge of the component’s (e.g., a transformer’s) design, operating (usage) characteristics and even past history of failures and repairs.

- **fault repairs**: identifying the repair actions that need to be carried out on the faulty component (e.g., a transformer) to mitigate the fault; then testing and ensuring that the repair actions have indeed eliminated the fault. Like fault diagnosis, fault repairing is also a very technical activity that needs detailed knowledge of the design of the component. Very often, cost and time considerations are important in selecting the appropriate repair actions. If a component requires too much time or efforts for its repairs, then one may decide to simply replace the component with a new (or normal) one and repair the faulty component later in the workshop.

There are other aspects of the active maintenance process that are important in real life. A power distribution network may sometimes suffer from multiple faults, a common enough situation, particularly in developing countries. In that case, one may need to prioritize the faults in terms of their urgency and carry out the above activities for each fault in order of its priority.

The fault prioritization may take into account factors such as the number of people affected, the nature of services and installations affected (e.g., a fault that affects important government institutions, hospitals or factories may need to be given higher priority) and so forth.

Further, in case of multiple faults, one may need to schedule the resources and people appropriately so as to minimize the overall time needed to repair all faults. This activity would take into account the time taken to travel up to the fault location, the nature of the fault, and the estimated time required to repair the fault. Standard operations scheduling techniques (e.g., linear and integer programming) are often used to minimize the fault service time.

Finally, the maintenance process also involves **planned maintenance**, in which one may attempt to predict the components that are likely to fail next and perform preventive maintenance so as to minimize the number of faults that may occur (as well as to minimize the mean time between failures) in future. A large number of forecasting and optimization techniques may be brought to bear on the various sub-problems involved in planned maintenance.

Currently, the entire maintenance process for power distribution networks is in various stages of automation in developing countries including India. In this paper, we propose an automated technique to handle the fault localization activity. In this activity, given a fault (e.g., a section of the power distribution network that has lost the power), one needs to identify the **root cause** of the fault i.e., the primary faulty component(s) whose failure has led to the observed power outages. As mentioned earlier, the fault localization problem is solved automatically and efficiently if the power distribution network is equipped with SCADA-enabled fault detection equipment. Our technique applies only to those power distribution networks which are not equipped with such sophisticated failure detection equipment.

### 25. OPERATIVE DIAGNOSIS

Operative diagnosis is a recently proposed set of algorithmic techniques to perform automatic diagnosis for certain kinds of systems. The discussion of operative diagnosis in this section is based on [9], [2], [11]. In operative diagnosis, we represent a system as a system graph (or network) in which vertices represent system components and edges (or links) between two components represent the flow of fault from one component to another. **Alarms** are attached to some of the components. An alarm attached to a component **u rings** if the component **u** itself fails or if fault from another component reaches **u**, causing **u** also to fail; otherwise, the alarm at component **u** is **silent**. A is the set of components to which an alarm is attached. Given a system graph and a subset of alarm vertices which are ringing, the operative diagnosis techniques identify the root cause of the fault, which is a minimal set of components whose failure can precisely “explain” the ringing and non-ringing alarms.

**Definition 1.** A system graph is a directed $G = (V, E, A)$, where the vertices in $V$ correspond to components in the system, there is a directed edge $(u, v) \in E$ from component $u$ to component $v$ if a fault in $u$ can flow into $v$ (i.e., if $u$ becomes faulty then $v$ also eventually becomes faulty) and $A \subseteq V$ is a set of vertices to each of which an alarm is attached.

**Definition 2.** Let $G = (V, E, A)$ be a given system graph. Given a non-empty subset $R \subseteq A$ of ringing alarms, the **operative diagnosis problem** is to find a minimal set of components $D \subseteq V$ such that

1. for every ringing alarm $a \in R$, there is a path from some component in $D$ to $a$; and
2. for every silent alarm $s \in A – R$, there is no path from any component in $D$ to $s$.

Such a set $D$ of faulty components is called a **diagnosis**. A diagnosis $D$ is minimal if no proper subset of $D$ is itself a diagnosis.

Thus given a system graph, a set of ringing alarms and a set of silent alarms, a diagnosis is a set of components that “explains” the ringing and silent alarms in the precise sense as defined above. Note that operative diagnosis does not identify what is actually wrong with a faulty component: it merely identifies the faulty components from the flow of the faults in the system. For example, it may identify a transformer as a faulty component, but...
it will not actually identify the reason for the failure of the transformer (e.g., burnt winding). In this sense, operative diagnosis really performs fault localization and not fault diagnosis.

An operative diagnosis algorithm takes a system graph $G$ and sets $R$ and $S$ of ringing and silent alarms as input ($A = R \cup S$) and produces a collection of diagnosis $D$ as output. In general, many different sets of components can explain the given ringing and silent alarms i.e., a diagnosis is not unique. Hence, a diagnosis algorithm may be designed to produce all possible diagnosis sets. These diagnoses can then be ordered (or ranked) in terms of various criteria. For example, one may prefer a smaller diagnosis over a larger diagnosis, due to the reason that single component failure is much more likely than say a failure of 3 components. If appropriate history is available, then one may use the (estimated) probability of failure of various types of components to rank the collection of diagnoses. Assuming component fail independently of each other, the probability of a failure of a $k$-component diagnosis $D = \{u_1, u_2, \ldots, u_k\}$ is the product of the probability of failure $P(u_i)$ of the individual components: $P(D) = P(u_1) \cdot P(u_2) \cdot \ldots \cdot P(u_k)$. The diagnoses $D_1$, $D_2$, ..., can then be ranked in the descending order of their probabilities of failure $P(D)$.

To simplify matters, a diagnosis algorithm may sometimes be designed produce a set of potential faulty sources, denoted $P$, which is a union of all diagnoses. One or more components in $P$ are guaranteed to be faulty.

When no limit is assumed on the maximum number of faulty components that can be present, the problem is NP-complete [9]. Hence, to simplify the problem, one can attempt to find a diagnosis, assuming that at most $k \geq 1$ components are faulty at any given time. When $k = 1$ or $k = 2$ this is called single fault and double fault assumption respectively.

Fig. 1(a) shows a system graph, a set of ringing alarms $A_R = \{7, 8, 13\}$ and a set of silent alarms $S = \{6, 9, 10\}$. Under the single fault assumption, $D = \{3\}$ is the only possible diagnosis. Fig. 1(b) shows another system graph with the set of ringing alarms $A_R = \{6, 7\}$ and set of silent alarms $\{8\}$. Under the single fault assumption, there is only one diagnosis $D = \{3\}$. However, under the double fault assumption, there are several diagnoses: $\{2, 3\}$, $\{2, 7\}$, $\{3, 6\}$, $\{3, 7\}$ and $\{6, 7\}$. In this case, instead of generating these $5$ sets, one may generate the set of potentially faulty sources $P = \{2, 3, 6, 7\}$. Two components in $P$ are definitely faulty. One may perform additional tests to find out which two components in $P$ are actually faulty.

It is tacitly assumed in the above discussion that there is zero (or negligible) delay between a component becoming faulty and the propagation of this fault to other nodes and ringing of alarms. Such systems are called zero-time systems. Electrical power distribution networks may be considered as zero-time systems, since the time required for the flow of fault (power outage) between components is often very short and negligible. In practice, there are non-zero time systems, for which there may be a finite delay between the occurrence of a fault and its propagation to other nodes. Delay in non-zero time systems can be known and fixed or unknown and time varying. Different diagnosis algorithms are needed for non-zero time systems.

Suppose we assume that at most one component is faulty. In that case, a simple algorithm called FORWARD [9] can detect the collection of all singleton diagnoses. Let $A_R \subseteq A$ denote the set of all ringing alarms. Let $B_R \subseteq A_R$ denote the largest subset of $A_R$ such that no vertex of $B_R$ is reachable from any other node of $A_R$. In Fig. 1, $A_R = \{7, 8, 13\}$ and $B_R = \{7, 8\}$. For a vertex $v$, the restricted to-set of $v$, denoted $RTS(v)$, is the set of all vertices $u$ such that the there is a path from $u$ to $v$ not containing any alarm, except possibly $v$ itself. In Fig. 1, $RTS(13) = \{11, 12, 13\}$. The FORWARD algorithm systematically examines each vertex $v \in RTS(v)$, for every ringing alarm $v \in B_R$, and examines whether $v$ correctly explains all alarms (both ringing and non-ringing) using say depth-first search. If yes, then $\{u\}$ is a possible diagnosis. In Fig. 1, $RTS(7) = \{1, 2, 3, 7\}$, $RTS(8) = \{1, 3, 4, 8\}$. Clearly, $\{3\}$ is the only diagnosis under the single fault assumption. [9] has proposed another algorithm called BACKWARD that is more efficient, though slightly more complex, than FORWARD under the single fault assumption.

The algorithm FORWARD can be easily extended to detect all 2-element diagnoses, under the double fault assumption (at most 2 components are faulty) as follows. Let $X$ denote the union of all sets $RTS(v)$ for vertex $v \in B_R$. Systematically examine every pair of vertices $x, y \in X \times X$ and check whether $\{x, y\}$ “explains” the ringing alarms. More precisely, start two depth-first searches from $x$ and $y$ and store all the alarms visited during these searches. If the union of these two sets exactly equals $A_R$, then $\{x, y\}$ is a possible diagnosis. Ref. [2] has proposed a more efficient algorithm called D-FAULTS for the double fault assumption.

The general multiple faults operative diagnosis problem, where there is no assumption about how many components may have failed, has been shown to be NP-COMPLETE [9]. See [11] for an efficient heuristic algorithm for the multiple faults operative diagnosis problem.

26. OPERATIVE DIAGNOSIS OF ELECTRICAL POWER DISTRIBUTION NETWORKS

In this section, we apply the operative diagnosis framework to the problem of fault localization in electrical power distribution network. In this paper, we have used radial power distribution networks for illustration, although the same techniques apply to inter-connected networks as well. Consider the radial electrical power distribution network in Fig. 2, which shows interconnections between the 18 components such as substation (SS), transformers (DT), poles (P) etc., including 16 customers. Note that this radial network is a tree. The network can be considered also as a system graph. The vertices are the components in the distribution network and links between two components indicate an electrical connection (power flow) between them. The flow of power is affected by failure of the components; e.g., if component DT12 fails, then none of the customers C1 to C5 will have power. Thus the presence of a fault is indicated by the absence of power.

As stated earlier, we do not assume the availability of any sensors, which will indicate the points in the network which have lost power. There are no automated alarms in this network. But we can consider customer complaints as manual alarms. Whenever the electric supply to a region fails, it might be expected that a few of the affected customers will complain to the power distribution company. Each such complaint can be considered as an alarm. Thus in this case complaint from a customer can be treated as an...
alarm. However, a customer is a may-ring alarm, since not every customer may complain even when experiencing a power failure.

Thus we have a situation in which may-ring alarms, which are silent, do not necessarily indicate absence of faults at the corresponding nodes. In other words, we need a minimal set of potential failure sources such that every ringing alarm is reachable from it and a minimal number of silent may-ring alarms are reachable from it.

The operative diagnosis problem for an electrical network is as follows. Given an electrical power distribution network and a set of customers that have complained about not having electrical power, identify a collection of diagnoses where each diagnosis is a minimal set of potentially faulty components.

The operative diagnosis framework is particularly useful in the scenario where customer complaints are received and processed centrally in either a call centre, a consumer service centre or a central control room. Suppose at any particular time instant, the central office has received $n$ complaints from $n$ consumers. The number of faults that have caused these complaints is unknown.

An important first question is: can we automatically determine the number $k$ of faults for a given set of complaints? Then the operative diagnosis techniques can be used to determine all diagnoses of size $k$. For example, suppose that in Figure 2 consumers C2, C5, C14 and C16 have complained about loss of power. If $k = 1$ then the diagnosis will be \{CB1\} or \{SS1\}. However, if $k = 2$ then the diagnoses could be \{DT12, P24\}, \{CB12, P24\} etc. which are more sensible, since poles and transformers are much more likely to fail than sub-stations. In general, the “lower” level components (i.e., those nearer to the consumers) are more likely to fail than “higher” level components. So how do we decide whether $k = 1$ or $k = 2$ (or even 3 or 4) for these 4 complaints?

Clearly, the size $k$ of a diagnosis (for a given set of complaints) and the total distance of the faulty components from the consumers are related. “Climbing away” from the consumers decreases the number $k$ of faulty components needed to explain the given set of complaints. One possible way is to identify that optimal value of $k$ which balances these two conflicting needs (increase $k$ versus climb away from consumers). However, we have chosen another way, which makes more sense in the assumed “call centre” like scenario and which makes use of additional data about the alarms (their physical locations).

We reformulate the earlier question as follows. How do we decide whether all the 4 complaints should be regarded as symptoms of a single (common) set of $k$ faulty components, or whether we should group the complaints in some way and treat each group as symptoms of a specific failure? The motivation for this question is as follows. Suppose we know that consumers C1 to C5, C6 to C10 and C11 to C16 belong to 3 well-separated towns. Then based on the considerations of physical proximity, the complaints C2, C5 belong to one group and C14, C16 belong to another group. It is much more likely that each group of complaints is caused independently by a separate fault. Treating the complaints together would then lead to a wrong diagnosis. The idea is to group (cluster) the complaints according to their physical proximity and independently diagnose each group by assuming either $k = 1$ or $k = 2$, say. It is of course possible (though less likely) that one of the groups may actually be caused by say $k = 5$ faults.

Assuming that the locations of the alarms (customers) are known (e.g., in a GIS system), we could use a suitable spatial clustering algorithm such as DBSCAN [4] (see [6] for a comprehensive...
survey of clustering algorithms) to automatically identify clusters (groups) of complaints. Complaints within a group should be close to each other i.e., they should be from the same geographical area and complaints in different groups should be far away from each other. Alternatively, the grouping can be done manually. For example, if the GUI indicates the location of each complaint by a red dot then the operator can visually partition the complaints (e.g., by circling groups of red dots) red and identify the clusters. Once the clusters are identified, we can apply the operative diagnosis algorithm discussed above to each cluster and then perform fault localization.

Note that a “sufficient” number of complaints need to be present before the correct diagnosis is possible. For example, in Figure 2, for the component DT21 to be correctly identified (under the single fault assumption), one complaint from \{C6, C7, C8\} and one from \{C9, C10\} must be received.

We have implemented a prototype of such a fault localization system for an electrical power distribution network. The clustering of complaints is done manually. Each cluster of complaints is given as input to the system and the system outputs the collection of possible diagnoses (each diagnosis is a set of faulty components). We have also implemented a rule-based system for prioritization of the selected faults, so that repairs can be carried out in order of the fault priority.

### 27. RELATED WORK
Fault diagnosis in electrical power distribution networks is a well-studied area. Work in [14] reports a multi-agent approach in a scenario very similar to the one we have assumed (large network, little telemetry data). Each of the 5 agents performs a specific task: report fault symptoms; diagnose the location and type (permanent or transient) of fault on (a) high voltage circuits (b) sub-stations (c) low voltage networks; make final diagnosis.

Unlike our approach, they work with the input that consists of full telemetry data, such as stats of switches and relays etc. [13] describes a comprehensive health management system for power distribution systems. It includes a range of data analysis and decision making techniques (e.g., statistical analysis, data visualization, principal components analysis and multiple discriminant analysis) for performing several tasks. The focus is more on fault diagnosis rather than fault localization; e.g., to decide whether a given fault is caused by a tree or an animal.

Similar work is reported in [3]. [12] reports a system called SHERLOCK that diagnoses faults in power distribution networks. Each type of fault (e.g., feeder faults, busbar faults and transformer faults) is represented by an explicitly defined set of Prolog rules. Given the telemetry status data (e.g., circuit breakers) and the network structure, the given network is searched to identify possible hypotheses that explain the observed symptoms and given rules. A single fault assumption is made. [5] present an approach where a distribution network is represented as a Petri-net and a simple procedure is presented to diagnose system faults based on telemetry status data for relays and circuit-breakers. Other work related to fault diagnosis in power distribution systems is presented in [7] and [10].

We have already reviewed basics of the operative diagnosis framework; for further details, please refer to [9], [2], [11]. A large number of different approaches have been proposed for fault diagnosis in different contexts: rule-based, graph-theoretical (other than operative diagnosis), fault-tree based, model-based, based on statistical techniques such as estimation and so on. Please see [1] [8] for a review.

![Figure 2. A radial electrical power distribution network.](image)
operative diagnosis framework can be applied to other activities in
the maintenance of power distribution networks.

29. ACKNOWLEDGMENTS
I thank Prof. Mathai Joseph for his support during this work. Thanks to Somnath Byabotta, Vaishali Kulkarni and other colleagues in TCS for their help. Thanks to the reviewers, whose useful comments helped me in improving the paper. I appreciate the encouragement from Dr. Manasee G. Palshikar.

30. REFERENCES


