A Bayesian Approach to Fault Isolation with Application to Diesel Engine Diagnosis

Anna Pernestål and **Mattias Nyberg**
Scania CV AB
Sweden
{anna.pernestal, mattias.nyberg}@scania.com

Bo Wahlberg
KTH Signals, Sensors and Systems
Sweden
bo@s3.kth.com

Abstract

This paper considers a Bayesian approach to fault isolation. Given a set of measurements from the system, and a set of possible faults, the task is to calculate the probability that the faults are present. This probability can then be used to rank the faults, or for decisions on fault sccomodation. The method requires the conditional probability distribution describing how the measurements react to the faults. In particular, the structure of dependencies between the tests is important. Knowing the structure facilitates efficient computation methods and makes it possible to reduce the memory capacity needed. In this paper, the structure is estimated from training data using Bayesian methods. The method is applied to diagnosis of the gas flow in a diesel engine.

1 Introduction

Fault isolation concerns the problem of localizing faults in technical processes. This is a most important problem in all field of industrial systems. Our motivating application is on-board fault isolation for diesel engines, where maintenance and repair procedures together with new emission regulations put challenging demands on the corresponding diagnosis systems. Other challenges are noise and model errors, which introduces uncertainty to the diagnosis process, and the limited storage capacity in the on-board control unit where the isolation system should be implemented. Further, industrial system are often large and complex, and it is impossible to build a complete model that is executable in the on-board control unit for the whole system.

The diagnosis system is structured as in Figure 1. This architecture is commonly used for diagnosis in the FDI community, for example when utilizing structured residuals, see [Gertler, 1998]. Further, it is one of the architectures used in industrial applications, and with this motivation we will use it in the present work.

The process to be diagnosed is assumed to consist of a set of components, which can be faulty or non-faulty. The components are monitored by precompiled diagnostic tests. An example of a diagnostic test is a thresholded residual. In the isolation system, the outputs from the tests are used to make inference about possible present faults.

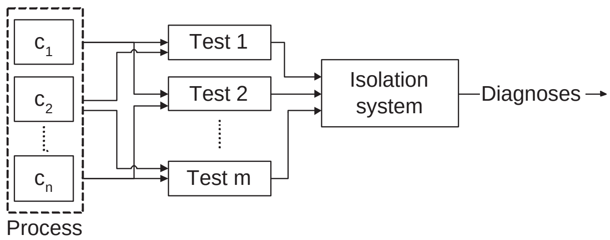

Figure 1: An example of relations between components c_i, tests and the isolation system.

In this work, the diagnostic tests are assumed to be given, and we will focus on the isolation system in Figure 1. The isolation system computes diagnoses, i.e. the combinations of faults that can explain the outputs from the tests. With the test results as our observations, this is the same definition of diagnoses as in [de Kleer *et al.*, 1992]. One problem is that already for small sized processes there can be many diagnoses, and hence a main requirement on the isolation system is that the diagnoses should be ranked after how probable they are. A second requirement is set by the limited processor and memory storage capacity in the on-board control unit.

In this work, a Bayesian approach is used for fault isolation. Given a set of test results, and a set of possible faults, the probabilities that different faults are present are computed. These probabilities are called posterior probabilities and can be used to rank the faults, or for decision making about fault accommodation. In order to compute the posterior, the conditional probability describing how the tests react on the faults is needed. In particular, the parameters and the structure of dependencies in the conditional distribution are needed. A complex structure, allowing a lot of dependencies, will increase the storage capacity needed. On the other hand, a too simple structure will affect the performance of the isolation.

There are two key contributions in this work. The first is that the structure of conditional probability is used as a design variable in the construction of the isolation system. The second is that Bayesian methods are used for estimation of the structure and the parameters of the conditional probability distribution from training data. Here, training data is the outputs from the diagnostic tests under different working conditions.

The main advantage of estimating the structure from training data is that no explicit knowledge about the process is needed. This is an advantage since in many industrial applications the system to be diagnosed is large and complex, and it is impossible to build a complete model of the whole system.

When the structure of dependencies and the parameters of the conditional probability is known, a Bayesian network can be set up and computationally efficient methods for probabilistic inference can be used, see for example [Lerner, 2002], [Lu and Przytula, 2005] or [Jensen, 2001].

2 Related work

When diagnosing complex systems, model errors, noise, and disturbances introduce uncertainties in the diagnosis computation. Several methods that handle the uncertainty have been proposed in the literature. In [Colin N. Jones and Lawrence, 2002] the PGDE (Probabilistic General Diagnostic Engine) algorithm is presented. In the PGDE the logic reasoning used in [Reiter, 1992] is combined with a measure of the belief in different diagnosis. In [Touaf and Ploix, 2004] the problem of uncertainty is solved using fuzzy logic methods. In [Pulido *et al.*, 2005] several isolation algorithms are combined, and the resulting algorithm is applied to uncertain models. In the present work, probabilistic reasoning is used, to compute the probability for different diagnoses.

Other probabilistic methods can be found in the literature. The Sherlock algorithm [de Kleer and Williams, 1992], as well as its precursor the GDE (General Diagnostic Engine) [de Kleer and Williams, 1987], contains a part, that corresponds to our isolation system and where probabilities for the diagnoses are computed. Those algorithms are designed for systems without noise, and the conditional probability distributions are assigned constant values, depending only on whether the measurement from the system is consistent with a fault or not. In the present work, training data is used to estimate the conditional probability distributions. If there is no training data available for the estimation of the underlying probability distributions, our algorithm is basically the same as in Sherlock and GDE.

In [Lerner *et al.*, 2000], [Schwall and Gerdes, 2002] and [Lu and Przytula, 2005] probabilistic reasoning for isolation is successfully used on noisy systems, utilizing Bayesian networks. In these three works knowledge about the process to be diagnosed is required to set up the structure for the probabilistic reasoning. In [Lerner *et al.*, 2000] the model of the system is translated into a Temporal Casual Graph and the structure of a Bayesian network is learned from it. In [Schwall and Gerdes, 2002] the structure of the model is given as input to the design of the isolation. In [Lu and Przytula, 2005] a known structure is used, and focus is on effective methods for solving the inference.

3 Problem Formulation

This work consists of two separate problems. The first is how to estimate the structure of dependencies and relations between the tests and the components, given a set of training data. This part is referred to as the *structure problem*. The second problem is to utilize the structure to compute the diagnoses when test results arrive to the isolation system, and is called the *isolation problem*. The structure problem requires tedious computations, but can be performed once and off-line. The isolation problem is performed on-line, where the computational and storage capacity is limited.

The process to be diagnosed consists of a set of N_C components, which can be faulty or not faulty. Enumerate the components, and let c_i be a variable with domain $\{NF, F\}$, where F means that the i:th component is faulty and NF that it is not faulty. The variable c_i is called the *behavioral mode* of the i:th component, [de Kleer *et al.*, 1992].

Assume that there exist N_D diagnostic tests. The diagnostic tests can be discrete or continuous. Here we will assume that the tests are discrete, because this is actually the case in our application, and because it simplifies the presentation. Note however that this is not necessary for the methods in general.

The diagnostic tests are assumed to be given, but we require no knowledge about their explicit construction. The only information needed is which faults that can possibly affect each. If a test is affected by a certain fault, it is said to be able to detect the fault, but due to model errors and noise it does not necessarily detect it. Enumerate the tests and let d_i denote the test result from test i. For example for a binary test, the test result can be either 1, indicating that a fault is detected, or 0 if no fault is detected.

The prior knowledge about the relations between tests and components can be presented as an isolation structure where an X at position (i, j) means that test i can react to a fault in the j:th component, but it does not necessary react every time the fault is present. A 0 at position (i, j) means that test i and component j are not related. For example, with three components and three tests, the isolation structure can look like

$$\begin{array}{c|ccc} & c_1 & c_2 & c_3 \\ \hline d_1 & X & 0 & X \\ d_2 & X & X & 0 \\ d_3 & 0 & X & X \end{array} \tag{1}$$

Let $C = [c_1, \ldots, c_{N_C}]$ be an assignment of behavioral modes to all components in the system, and let $D^t = [d_1, \ldots, d_{N_D}]$ be the test results at time t. We call C the system behavioral mode. The isolation problem is to compute the probability for different assignments of behavioral modes to all components in the systems, given the test results a certain time t,

$$P(C|D^t), \tag{2}$$

also referred to as the *posterior* probability. The probability (2), as well as all other probabilities in the following, should also be conditioned on the prior knowledge about the process, but for notational convenience we leave this out unless it is especially important. In the following we consider only the test results from a certain time, and we will suppress the index t to simplify notations.

To compute (2), use Bayes' rule,

$$P(C|D) = \frac{P(D|C)P(C)}{P(D)}, \tag{3}$$

where $P(C)$ is the prior probability for the system behavioral mode. These priors are assumed to be known. In real systems this represents the knowledge of the quality of the components. In (3) the denominator $P(D)$ is a normalization factor, which can be computed using marginalization over all possible system behavioral modes,

$$P(D) = \sum_{C} P(D|C)P(C). \tag{4}$$

The probability distribution $P(D|C)$ is called the *likelihood* for C, and it will now be shown how it can be estimated from data.

For the on-line isolation, the likelihood is stored as a table, and when test results arrive to the isolation system, the probabilities for different values of C given the data D are computed using (3). The table can be very large, for example assuming binary tests the number of elements needed for storage is $2^{N_C+N_D}$. Even for a small process containing only ten tests and ten components this table has more than a million elements, and the storage of the table is infeasible. To reduce the storage capacity needed for the likelihood $P(D|C)$, it can be factorized into mutually independent factors.

One naive approach is to assume that all tests are independent given the behavioral modes. This assumption is generally not true. Examples on situations where tests are dependent are when several faults have the same root cause, when one test can cause another test to react, when there are errors in the underlying models for the diagnostic tests, and when the probability that a test reacts is dependent on the working point or the environment. Measurements of outputs from the tests in engine diagnosis have shown that some tests certainly are dependent, while others are independent. Thus, assuming that all tests are independent, the posterior probabilities for the system behavioral mode will be incorrect.

Instead, partition the tests into M subsets, such that the tests in different subsets are mutually independent, or can be assumed to be mutually independent, but the tests in the same subset can be dependent. Let the maximum numbers of tests in a subset be L. Let I_i be an index vector, containing the indices for the tests that is in subset i, and $D[I_i]$ be the tests with indices in I_i. Then the partition of the tests D into M subsets gives the factorization

$$P(D|C) = P(D[I_1]|C)P(D[I_2]|C)\ldots P(D[I_M]|C) \tag{5}$$

of the likelihood. Each of the distributions $P(D[I_i]|C)$ can be represented by a table, which maximum size is determined by L. To avoid too large tables, and decrease the storage capacity needed, limits on L is used. For the case where the tests are binary, the maximum number of elements in each subset is 2^{N_C+L}. Although one table for each factor is needed, the total storage capacity required is reduced.

In (5) the subsets $D[I_i]$ are of different and unknown size. Also, the number of factors, M, in the factorization is unknown. Besides the factorization of the likelihood, the parameters of the distributions $P(D[I_i]|C])$ must be estimated. For the estimation, assume that we have a set of training data $\mathscr{D} = [\mathscr{D}_C \ \mathscr{D}_D]$, where $\mathscr{D}_C$ are the behavioral modes $\mathscr{D}_D$ the corresponding test results. The structure problem can now be stated as estimating the factorization (5) and the parameters of the underlying structure, given the training data $\mathscr{D}$. The structure problem can also be thought of in terms of a model selection problem. Here, we assign a model of the class of two-layer Bayesian networks, and use training data to estimate the best structure.

4 The Structure Problem

The structure problem can be visualized graphically as going from the left graph to the right graph in Figure 2. The left graph represents our prior knowledge about the relations between tests and components (solid lines), and the unknown relations between tests (dashed lines). The right graph represents the estimated structure, where tests that are dependent are grouped into the same node. In Figure 2 the left graph represents the structure given by (1), where an X at position (i,j) in (1) gives a solid line between test i component j. The right graph is an example where the tests one and three are grouped.

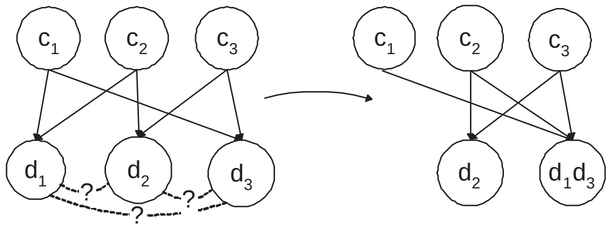

Figure 2: The structure problem can be represented as going from the left graph to the right.

To estimate the structure from data, a measure of how well a structure fits the training data is needed. In [Wolf, 1995] the χ^2-test is compared with a Bayesian approach. The disadvantage with the χ^2-test is that it is accurate only for large data sets. With the Bayesian approach, the probability that a certain structure is the underlying structure, given the training data, is computed. This is valid also for small sets of data, see [Jaynes, 2001] and [Wolf, 1995]. The Bayesian approach suits this problem, since the training data consists of few examples for system behavioral modes which are unlikely too occur. In the Bayesian approach prior probabilities for the different structures must be given. In this work an uninformative prior, ranking all structures as equally likely, will be used [Wolf, 1995].

To keep notations simple, the method will be illustrated with a simple example, but it is straight forward to generalize all reasoning to larger problems. In the example, there are three components, represented by the variables c_1, c_2 and c_3, three tests, and the maximum factor size L is set to 2. The relations between tests and components is given by the isolation structure (1). First, the structure is estimated, and then, given the structure, the parameters in the distribution is estimated.

4.1 Structure estimation

We search a factorization (5), or in other words the index sets I_i, $i = 1, \ldots, M$, that suits all different assignments of system behavioral modes C. To achieve this, we assume that the index sets in (5) are the same as in

$$P(D) = P(D[I_1])P(D[I_2])\ldots P(D[I_M]). \tag{6}$$

Note that it is only the structure, i.e. the index sets I_i, $i = 1\dots M$ and the number of elements M that are assumed to be the same in (6) and (5), and not the probabilities themselves.

This assumption is reasonable, since if two subsets $D[I_j]$ and $D[I_k]$ are independent, the knowledge of C will not make them dependent. On the other hand, if $D[I_j]$ and $D[I_k]$ are dependent, the knowledge of C can make them independent. As will be shown in Section 7.1, this will not affect the isolation performance, but only increase the storage capacity needed to perform the on-board isolation.

To bias the factorization such that it is better suited for more important faults, more training data from those behavioral modes can be used. Here, all faults are assumed to be equally important and equal amount of training data is used from each system behavioral mode.

Now, we introduce some notations. The distributions can be represented by multidimensional arrays, with one dimension for each test. Let $p = P(D)$, and for the marginal distributions $p^{12} = P(d_1,d_2) = \sum_{d_3} P(d_1,d_2,d_3)$, where we sum over all possible values of d_3 etc. Let l_r be the number of elements in p^r, $r = 1\dots3$. For the elements in the distributions, let $p_{ijk} = P(d_1 = i, d_2 = j, d_3 = k)$, $p^{12}_{ij} = P(d_1 = i, d_2 = j)$ and so on. Correspondingly for the data, let n_{ijk} be the number of observations with $d_1 = i, d_2 = j, d_3 = k$, and $n^{12}_{ij} = \sum_{k=1}^{l_3} n_{ijk}$ etc. The total amount of data is $N = \sum_{i,j,k=1}^{l_1,l_2,l_3} n_{ijk}$.

In our example p can be factorized in four different ways: such that all tests are independent, or such that two tests are dependent while the third is independent of the other two. In the example, use H_0 to denote the hypothesis "all three variables are independent" and H_q, $q = 1\dots3$ to denote "variable q is independent and the other two are dependent". Given a hypothesis H_q, $q = 0\dots3$, p can be factorized in a certain way. For example H_1 means that $p = p^1 p^{23}$. We search the probabilities for the different factorizations (6) given the training data $\mathscr{D}$. Since we only want one structure, we use the Maximum a posteriori (MAP) estimate $H^* = max_q P(H_q|\mathscr{D})$. Bayes' rule gives

$$P(H_q|\mathscr{D}) = \frac{P(\mathscr{D}|H_q)P(H_q)}{P(\mathscr{D})}, \tag{7}$$

where $P(\mathscr{D})$ is a normalization factor, which can be computed using marginalization,

$$P(\mathscr{D}) = \sum_{q=0}^{3} P(\mathscr{D}|H_q)P(H_q). \tag{8}$$

Here $P(H_q)$ is the prior probability for the different factorizations. We apply a prior that is zero for all partitions containing subsets with more than l elements, and constant for all other. The distribution $P(\mathscr{D}|H_q)$ can be computed using marginalization over all possible distributions,

$$P(\mathscr{D}|H_q) = \int P(\mathscr{D}|p,H_q) f(p|H_q) dp, \tag{9}$$

where $f(p|H_q)$ is the continuous distribution for p and $P(\mathscr{D}|p,H_q)$ is a multinomial distribution given by

$$P(\mathscr{D}|p,H_0) = \frac{N!}{\prod_{i,j,k=1}^{N} n_{ijk}!} \prod_{i,j,k=1}^{N} (p^1_i p^2_j p^3_k)^{n_{ijk}} \tag{10a}$$

$$P(\mathscr{D}|p,H_1) = \frac{N!}{\prod_{i,j,k=1}^{N} n_{ijk}!} \prod_{i,j,k=1}^{N} (p^1_i p^{23}_{jk})^{n_{ijk}}, \tag{10b}$$

and similarly for $q = 2,3$. The elements in distributions must be between 0 and 1, and for each distribution they must sum to one. The first criteria is regulated by the integration limits in (9). The latter criteria means that f is proportional to delta functions as

$$f(p|H_0) = f(p^1 p^2 p^3|H_0) \tag{11a}$$
$$\propto \delta(\sum_{i=1}^{l_1} p^1_i - 1)\delta(\sum_{j=1}^{l_2} p^2_j - 1)\delta(\sum_{k=0}^{l_3} p^3_j - 1)$$

$$f(p|H_1) = f(p^1 p^{23}|H_0) = \tag{11b}$$
$$\propto \delta(\sum_{i=1}^{l_1} p^1_i - 1)\delta(\sum_{j,k=1}^{l_2,l_3} p^{23}_{jk} - 1),$$

and similar for H_2 and H_3.

The integral (9) can now be solved using convolution and Laplace transform techniques [Wolf, 1995]. The result is

$$\int P(\mathscr{D}|p,H_0) f(p|H_0) dp = \frac{N!}{\prod n_{ijk}!} \Gamma(l_1)\Gamma(l_2)\Gamma(l_3) F_0 \tag{12a}$$

$$\int P(\mathscr{D}|p,H_1) f(p|H_1) dp = \frac{N!}{\prod n_{ijk}!} \Gamma(l_1)\Gamma(l_2 l_3) F_1 \tag{12b}$$

where $\Gamma(\cdot)$ is the gamma function and

$$F_0 = \frac{\prod_{q=1}^{3} \prod_{i=1}^{l_q} \Gamma(n^q_i + 1)}{\Gamma(N+l_1)\Gamma(N+l_2)\Gamma(N+l_3)} \tag{13a}$$

$$F_1 = \frac{\prod_{i=1}^{l_1} \Gamma(n^1_i + 1) \prod_{j,k=1}^{l_2,l_3} \Gamma(n^{23}_{jk} + 1)}{\Gamma(N+l_1)\Gamma(N+l_2 l_3)}. \tag{13b}$$

The expression (12) and (13) are similar for H_2 and H_3.

Now, we can compute $P(H_q|\mathscr{D})$ for all q. With the MAP estimate $H^* = \max_q P(H_q|\mathscr{D})$ for the structure, i.e. the index sets in (6) and hence also in (5), we can estimate the parameters in the factors in (5).

4.2 Parameter estimation

For the parameter estimation we use the same notation as in Section 4.1, but with subindex C to denote that we condition on the system behavioral mode, i.e. $p_C = P(D|C)$. We will use the MAP estimate p^*_C, that maximizes $f(p_C|\mathscr{D},H^*)$. Again, apply Bayes' rule,

$$f(p_C|\mathscr{D},H^*) = \frac{P(\mathscr{D}|p_C,H^*)P(p_C|H^*)}{f(\mathscr{D}|H^*)}, \tag{14}$$

where $P(p_C|H^*)$ is the prior for p_C given the partition H^*. In this work a prior that is uniform for all p_C which suits the partition H^* and the structure (1), and is zero for all other p_C's is applied. The denominator in (14) is a normalization factor, independent of p_C, and hence $P(p_C|\mathscr{D},H^*) \propto P(\mathscr{D}|p_C,H^*)$ for all p_C suitable to H^*. The MAP estimate is $p_C^* = \max_{p_C} P(p_C|\mathscr{D},H^*)$. From (5) and given H^* we know that we can factorize $p_C = \prod_{m=1}^M p_C^m$. The distribution $P(\mathscr{D}|p_C,H^*)$ is minimized under the constraint that all elements in each factor should sum to one, $\sum_i p_{C,i}^m = 1$, $m = 1 \dots M$, using Lagrange multipliers. The result is

$$p_{C,x}^{m*} = \frac{n_x^m}{N}, \tag{15}$$

for $x = 1 \dots l_m$. With $p_C^{m*} = [p_{C,1}^{m*} \dots p_{C,l_m}^{m*}]$ and $p_C^* = \prod_{m=1}^M p_C^{m*}$ we know, together with H^*, both the structure and the parameters in (5), and the isolation problem can be solved using probabilistic inference.

No training Data

If there is no training data available, other ways of assigning the the structure and the probabilities are needed. Using the principle of indifference [Jaynes, 2001], we assign the probabilities

$$p_C = \begin{cases} 0 & \text{if } D \text{ is inconsistent with } C, \\ 1 & \text{if } D \text{ is surely consistent with } C, \\ \frac{1}{K} & \text{otherwise.} \end{cases} \tag{16}$$

Here K is the number of values that D can take given C. In this case, the structure will not affect the result, and the assumption that all data is independent can be used. For example, using the isolation structure (1) and assuming binary tests, this gives $P(d_1 = 1|C = [NF,NF,F]) = \frac{1}{2}$ and $P(d_1 = 1|C = [NF,NF,NF]) = 0$. This is basically the same approach as used in [de Kleer and Williams, 1987] and [de Kleer and Williams, 1992].

5 The Isolation Problem

The on-line isolation is solved by computing the posterior probability for the system behavioral modes, given the test results at a certain time and using the structure H^*, the estimated likelihoods p_C^* and the information about the prior $P(C)$. Denoting the prior information with I, the posterior probability is

$$P(C|D,H^*,p_C^*,I). \tag{17}$$

To compute (17) efficiently, a Bayesian network can be set up, using the structure and the parameters learned from the structure problem. Standard algorithms for reasoning in Bayesian networks can be used, see [Jensen, 2001] or [Lerner, 2002] for examples. There are also algorithms for computing the k most likely explanations of the data, with even less complexity [Lerner, 2002].

6 Performance Measure

In the present paper, isolation systems that can be expressed by a two-layer Bayesian network is designed. By choosing different values of L, different isolation systems within this class is designed. Further, there are the two extreme cases, assuming that all tests are independent, i.e. $L = 1$, and using no assumptions on independence. Let $\mathbb{I}$ denote an isolation system. Then the output from the Bayesian isolation systems, the posterior, is $P(C|D,\mathbb{I})$.

In order to compare the performance of two isolation systems, a performance measure is needed. We suggest as an optimal isolation system, a system that gives the posterior probability one for the true underlying system behavioral mode. For probabilistic isolation systems, define the Expected probability of correctness,

Definition 1 (Expected probability of correctness) *Let D_{C^*} be data generated when the system behavioral mode C^* is present, and let $\mathbb{I}$ be a probabilistic isolation system. Then the expected probability of correctness is*

$$\mu(C^*,\mathbb{I}) = E\left\{P(C^*|D_{C^*},\mathbb{I})\right\}, \tag{18}$$

where the expectation is over data.

The measure μ gives the expected probability assigned to the system behavioral mode that is really present. The optimal value of μ is one. This measure gives one number for each system behavioral mode, which is interesting since the behavioral modes can be differently difficult to isolate.

To summarize the expected probability of correctness into one number, use the average over all system behavioral modes,

$$\bar{\mu}(\mathbb{I}) = \frac{1}{m}\sum_C \mu(C,\mathbb{I}). \tag{19}$$

Another measure that relates to the isolation system performance is the probability that a correct diagnose is done, if the system behavioral mode with largest posterior probability is chosen as the diagnosis. In other words, given data D_{C^*}, from the system behavioral mode C^*, what is the probability that $P(C^*|D)$ is the largest posterior probability? We call this measure the expected probability of correct classification and write it

$$\mu_{cc}(C^*,\mathbb{I}) = E\left\{P\left(\hat{C} = C^*\right)\right\}, \tag{20}$$

$$\text{where } \hat{C} = \max_C P(C|D_{C^*},\mathbb{I}) \tag{21}$$

and the expectation is over data. The optimal value of $\mu_{cc}(\mathbb{I})$ is one. Also this measure gives one number for each behavioral mode. To summarize μ_{cc}, use the average,

$$\bar{\mu}_{cc}(\mathbb{I}) = \frac{1}{m}\sum_C \mu_{cc}(C,\mathbb{I}). \tag{22}$$

Note that choosing the system behavioral mode with largest probability as the diagnosis is only one of the interpretations of the output from the isolation system. There are more clever ways to interpret the results. How to interpret the results from an probabilistic isolation system is further discussed in Section 7.2.

7 Diesel Engine Diagnosis

The Bayesian isolation approach is applied to the diagnosis of the gas flow of a diesel engine with EGR (Exhaust Gas Recirculation) and VGT (Variable Geometry Turbine). A

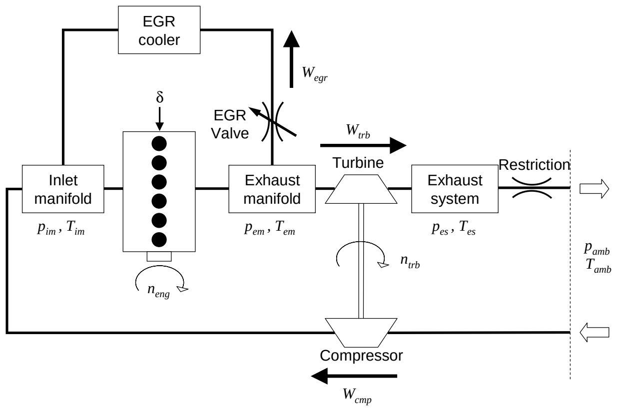

Figure 3: A schematic figure of the gas flow through the diesel engine with EGR and VGT

schematic figure of the gas flow is given in Figure 3. In the system there are ten components, to be diagnosed, listed in Table 1. In this example, all components considered are sensors, but other kinds of components, such as pipes, actuators etc. can be diagnosed with this method as well.

Table 1: The sensors in the engine system

p_{em}	exhaust gas pressure
p_{im}	inlet manifold gas pressure
T_{im}	inlet manifold temperature
p_{amb}	ambient pressure
T_{amb}	ambient temperature
u_{EGR}	EGR valve position
u_{vgt}	VGT valve position
w_{cmp}	flow through the compressor
n_{eng}	engine speed
n_{trb}	turbine speed

To make the results easier to overview, only the three components p_{em}, p_{im}, and n_{trb} are diagnosed in this example, while the other seven are assumed to function correctly. An extension to diagnosis of all sensors is straight forward. All possible combinations of faults of the three components are considered. This gives $m = 8$ possible system behavioral modes. For the system behavioral modes we use a short notation. For example to denote that components p_{em} and p_{im} are functioning correctly and component n_{trb} is faulty, we let $C = [p_{em}, p_{im}, n_{trb}] = [NF, NF, F]$ be represented by $C = [001]$.

There exists a complex model of the diesel engine process, from which about 60 residual generators can be found [Einarsson and Ahrrenius, 2004]. Due to limitations in the capacity of the on-board control unit, not all 60 residual generators can be executed. In this example, five of the 60 residuals are used. The residuals are thresholded, and the thresholded residuals are used as the diagnostic tests. Here, the tests are binary, but this is not a requirement for the method. The experiments are done on data collected from the engine in a real driving situation.

Four different isolation systems, are set up:

- $\mathbb{T}$ No assumption of independence
- $\mathbb{H}^*$ The most probable structure for some given requirements, H^*
- $\mathbb{N}$ The naive assumption that the tests are independent, $L = 1$
- $\mathbb{D}$ No training data

The diagnosis system $\mathbb{T}$ is of course infeasible when considering larger systems, but is used here because it uses all the information given by the training data, and gives in some sense the best possible structure. The system $\mathbb{D}$, designed without training data, is implemented according to (16). This turned out to perform very poor on the current example, and the result is only given in the summary of the experiments in Table 4.

For the design of the isolation system $\mathbb{H}^*$, the requirement $L = 2$ is used. This reduces the required storage capacity from $2^8 = 256$ for the system $\mathbb{T}$ to, in worst case, $2 \times 2^5 + 2^4 = 80$. In practice, the storage capacity needed will be even smaller, since the tests in each partition will not be related to all components.

7.1 Experimental Results

To design the isolation system $\mathbb{H}^*$, the probabilities for all possible structures with $L = 2$ were computed. The probabilities for the five most probable structures, normalized with the probability of the most probable structure, are given in Table 2. It is clear that the partition $[14,\ 23,\ 5]$ is far more probable than the other, and also that there are similarities between the most probable structures. Experiments are run over 10000 Monte Carlo simulations based on data from real driving situations.

Table 2: The five most probable partitions normalized with the probability of the most probable partition

Partition, H_i	$P(H_i \vert \mathscr{D})/P(H^* \vert \mathscr{D})$
$14,\ 23,\ 5$	1
$15,\ 23,\ 4$	0.21
$1,\ 23,\ 4,\ 5$	0.17
$14,\ 23,\ 5$	0.1
$1,\ 23,\ 45$	0.07

The prior probabilities for all three faults are assumed to be equal, $p(c_i) = 0.1,\ i = 1, 2, 3$, and, although not necessary, we assume that they break independently.

To compare the performance of the isolation systems, data sets from different system behavioral modes are applied to the systems, and the probabilities for different diagnoses are computed. In Figures 4, 5, and 6 the probability, and its variance, for three different system behavioral modes and the three isolation systems $\mathbb{T}$, $\mathbb{H}^*$, and $\mathbb{N}$ are shown. The true behavioral modes are $C = [010]$, $C = [110]$, and $C = [110]$ respectively. For the first two system behavioral modes, the isolation systems $\mathbb{T}$ and $\mathbb{H}^*$ assign largest probability to the correct system behavioral mode, while the the system $\mathbb{N}$ does not. For the third system behavioral mode, in Figure 6, all

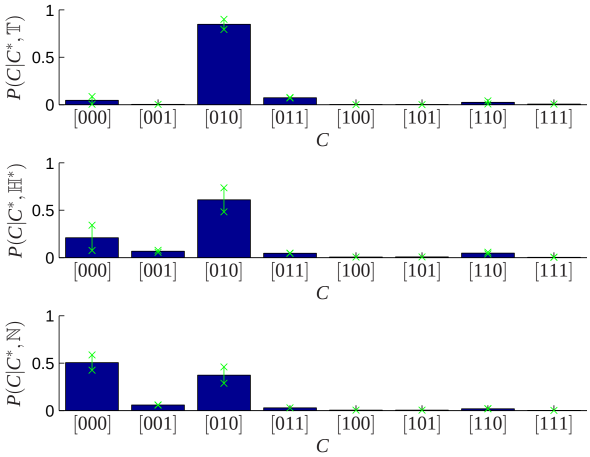

Figure 4: The average probability assigned to the different system behavioral modes for the isolation systems $\mathbb{T}$ (top), $\mathbb{H}^*$ (middle), and $\mathbb{N}$ (bottom). The lines show the variance. The true behavioral mode is $C^* = [010]$.

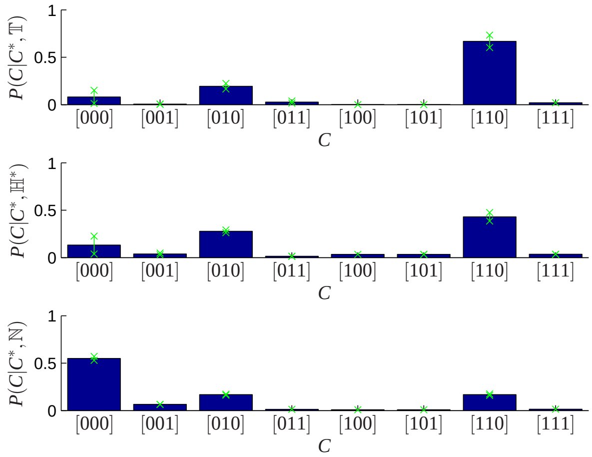

Figure 5: The average probability assigned to the different system behavioral modes for the isolation systems $\mathbb{T}$ (top), $\mathbb{H}^*$ (middle), and $\mathbb{N}$ (bottom). The lines show the variance. The true behavioral mode is $C^* = [110]$.

Table 3: The expected probability of correctness for three different system behavioral modes.

Isolation System	μ $C = [010]$	$C = [110]$	$C = [011]$
$\mathbb{T}$	0.84	0.69	0.11
$\mathbb{H}^*$	0.59	0.45	0.05
$\mathbb{N}$	0.30	0.18	0.02
$\mathbb{D}$	0.002	0.002	0.004

Table 4: The probability of correct classification and the average probability of correctness for all systems.

Isolation system	$\bar{\mu}_{cc}$	$\bar{\mu}$	storage needed
$\mathbb{T}$	0.44	0.40	256
$\mathbb{H}^*$	0.29	0.25	80
$\mathbb{N}$	0.18	0.18	80
$\mathbb{D}$	0.13	0.13	80

three isolation systems misses the underlying system behavioral mode.

The expected probability of correctness for the behavioral modes $C = [010]$, $C = [110]$, and $C = [011]$ are given in Table 3. All the values of μ are far from 1, even for the system $\mathbb{T}$, although no assumptions on independence are done in this system. The reason is that some system behavioral modes are difficult to isolate, for example $C = [011]$ shown in Figure 6. A numerical summation of all four isolation systems is given in Table 4. The values of $\bar{\mu}$ and $\bar{\mu}_{cc}$ of the isolation system $\mathbb{T}$ is largest, followed by the system designed with our method, $\mathbb{H}^*$.

The values of $\bar{\mu}_{cc}$ in Table 4 indicates that choosing the system behavioral mode with the largest probability as the diagnosis is not always a good way of interpreting the results. It is also interesting to note that the naive isolation system, and our designed isolation system needs the same amount of storage capacity for the likelihoods. This is not true in general, although L can often be chosen so that the storage capacity needed is significantly reduced compared to the system without restrictions.

The performance is very different for different system behavioral modes. In general multiple faults are more difficult to detect than single faults. The reason is that the priors for multiple faults are very small compared to the priors for single fault or no fault. One solution to overcome this problem is to consider data from several time steps. In this case the probability $P(C|D^t, D^{t+1}, \ldots D^{t+T}, \mathbb{I})$ for some $T > 0$ is used instead of the probability $P(C|D^t, \mathbb{I})$ as in the case above. This will decrease the influence of the prior, and increase the influence of the likelihood on the posterior. See for example [Jaynes, 2001].

7.2 Discussion

The experimental results show that isolation system based on the partitioned structure performs better than the isolation system based on the naive structure, but still it performs worse than the structure assuming no independences. One question is of course, how much performance can be gained using a larger L. The maximum L that can be used is given by restrictions on the memory capacity of the on-board control unit, but also a smaller L could perform sufficiently good. The accuracy needed is dependent on how the output from the isolation system is to be evaluated.

One way to interpret the output from a probabilistic isolation system is to use a cost function, and compute the expected cost of measures. The target is to minimize this expected cost.

So far we have focused on the storage needed to imple-

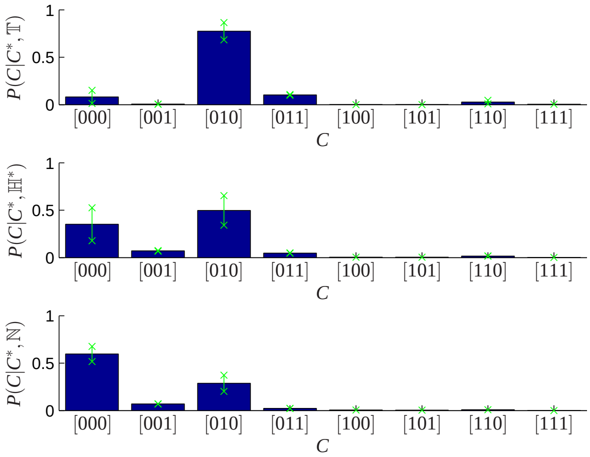

Figure 6: The average probability assigned to the different system behavioral modes for the isolation systems $\mathbb{T}$ (top), $\mathbb{H}^*$ (middle), and $\mathbb{N}$ (bottom). The lines show the variance. The true behavioral mode is $C^* = [011]$.

ment the isolation system, and seen that it is dependent on L. Also, the number of hypothesis H_q defined in Section 4 is intersting, since too many hypotheses can give numerical problems when solving the structure problem. The number of hypotheses increases with increasing L, and with increasing number of diagnostic tests. To extend this work to large scale problems, the increased search space for H_q^* must be handled. This extension is a challenge, but beyond the scope of this work.

8 Conclusion

In this paper Bayesian techniques for fault isolation is presented. The structures of the underlying conditional probabilities is used as a design variable, and they are estimated from training data. The Bayesian method was applied to diagnosis of the gas flow of a diesel engine.

Four different Bayesian isolation systems, with different degrees of dependence assumptions, were compared. The experiments was run on data from real driving situations. The result shows that if there is a dependence betweens tests, this dependence is important to take into account when designing the isolation system. The system designed with the new method performs best of the systems with the same order of complexity.

References

[Colin N. Jones and Lawrence, 2002] Gregory W. Bond Colin N. Jones and Peter D. Lawrence. Consistency-based fault isolation for uncertain systems with applications to quantitative dynamic models. In *DX 2002*, pages 36–42, 2002.

[de Kleer and Williams, 1987] Johan de Kleer and Brian C. Williams. Diagnosing multiple faults. *Artif. Intell.*, 32:97–130, 1987.

[de Kleer and Williams, 1992] Johan de Kleer and Brian C. Williams. Diagnosis with behavioral modes. In *Readings in model-based diagnosis*, pages 124–130, San Francisco, CA, USA, 1992. Morgan Kaufmann Publishers Inc.

[de Kleer *et al.*, 1992] Johan de Kleer, Alan K. Mackworth, and Raymond Reiter. Characterizing diagnoses and systems. *Artif. Intell.*, 56(2-3):197–222, 1992.

[Einarsson and Ahrrenius, 2004] Henrik Einarsson and Gustav Ahrrenius. Automatic design of diagnosis systems using consistency based residuals. Master's thesis, Uppsala University, 2004.

[Gertler, 1998] Janos J. Gertler. *Fault Detection and Diagnosis in Engineering Systems*. Marcel Decker, New York, 1998.

[Jaynes, 2001] B. T. Jaynes. *Probability Theory - the logic of science*. Camebridge University Press, Cambridge, 2001.

[Jensen, 2001] X. Jensen. *Bayesian networks*. Springer-Verlag, New York, 2001.

[Lerner *et al.*, 2000] Uri Lerner, Ronald Parr, Daphne Koller, and Gautam Biswas. Bayesian fault detection and diagnosis in dynamic systems. In *AAAI/IAAI*, pages 531–537, 2000.

[Lerner, 2002] Uri Lerner. *Hybrid Bayesian Networks For Reasoning About Complex Systems*. PhD thesis, Stanford University, Stanford University, October 2002.

[Lu and Przytula, 2005] Tsai-Ching Lu and K. Wojtek Przytula. Methodology and tools for rapid development of large bayesian networks. In *DX 2005*, pages 89–94, 2005.

[Pulido *et al.*, 2005] B. Pulido, V. Puig, T. Escobet, and J. Quevedo. A new fault localization algorithm that improves the integration between fault detection and localization in dynamic systems. In *DX 2005*, 2005.

[Reiter, 1992] Raymond Reiter. A theory of diagnosis from first principles. In *Readings in model-based diagnosis*, pages 29–48, San Francisco, CA, USA, 1992. Morgan Kaufmann Publishers Inc.

[Schwall and Gerdes, 2002] Matthew Schwall and Christian Gerdes. A probabilistic approach to residual processing for vehicle fault detection. In *Proceedings of the 2002 ACC*, pages 2552–2557, 2002.

[Touaf and Ploix, 2004] Samir Touaf and Stephane Ploix. Soundly managing uncertain decisions in diagnostic analysis. In *DX 2004*, 2004.

[Wolf, 1995] David Wolf. Mutual information as a bayesian measure of independence, 1995.