Enabling Reliable Detection of Failed Parts in Cyber-Physical Systems Using Unreliable Detection Sensors

Boris Kriheli[a,b] and Eugene Levner[a]
[a] Holon Institute of Technology, Holon, Israel, levner@hit.ac.il
[b]Ashkelon Academic College, Ashkelon, Israel, borisk@hit.ac.il

Abstract: Consider a problem of reliable detection of failures in Educational Cyber-Physical Systems (ECPS) consisting of several hundred computers, advanced audio and video devices and control mechanisms, as well as intelligent sensors, which are located at home and in the classrooms of the university; they can interact with each other via the Internet and serve for hybrid teaching and learning, for example, during and after the pandemic. Various probabilistic and statistical methods are known for assessing the probability of failure detection in the system, depending on the reliability of the imperfect sensors. Our approach is based on probabilistic safety assessment. To achieve the required level of safety, our key idea is to check each CPS component several times in succession. Using the formula for total probability and the Bayesian approach, we build a mathematical model for finding the minimum number of necessary repeated tests required for each component. We then develop an appropriate test scheduling algorithm and examine its complexity. Finally, we run computational experiments to detect failures in a real educational CPS serving for hybrid teaching; we compare the proposed algorithm with two known failure-detection algorithms and obtain encouraging practical results.

1. INTRODUCTION

Cyber-Physical Systems (CPS) are networks of physical and computer components that are interconnected to securely, flexibly, and efficiently manage integrated computing, networking, and physical processes. The CPS have pervaded everywhere; examples include the internet of medical robotic things (IOMRT), intelligent operating rooms and surgeries, computer-integrated manufacturing systems, robotics systems, hybrid educational systems, regional power grids, smart homes, etc. ([14]). Educational Cyber-Physical Systems (ECPS) consist of several hundred computers, advanced audio and video devices and control mechanisms, as well as intelligent sensors, that are located at home and in the classrooms of the university; they can interact with each other via the Internet and serve for hybrid online/face-to-face teaching and learning, for example, during and after the pandemic.

The reliability and safety aspects of such systems are crucial. Various probabilistic and statistical methods are known for assessing the probability of failure detection in the system, depending on the reliability of the imperfect sensors. Our approach is based on probabilistic safety assessment.

In this paper, we are looking at a multi-component ECPS that is known to have failed; it is necessary to determine what is the exact location of the failed component. For this, a system of unreliable sensors is used, which sequentially check the ECPS components. For any possibly failed part of the ECPS, the following data are known: (*a*) the cost and time of checking the component; (*b*) the initial probability that the component failed; (*c*) the probability of a false-negative test result for the sensors, that is, the probability that during the checks the unreliable sensors will not notice that a faulty component is out of order; (*d*) the probability of a false-positive test result, when the sensor erroneously classifies a healthy component as ”failed”; and (e) the required safety level, which is defined as a desired level of the probability of correctly detecting the failed ECPS component; this parameter is set by the decision maker in advance and far exceeds the known probabilities of obtaining correct results by individual sensors.

The discrete sequential failure search is an ancient operations research problem that has many civil and military applications (see, e.g., Matula [12], Levner [11], Kress et al. [7], Alpern et al. [2], Kriheli and Levner [8], Kadane [6], Kriheli et al. [9], Sotoudeh-Anvari et al. [15], and references therein). In this paper we develop a new algorithm that enables to efficiently solve a novel formulation of the problem.

To achieve the required level of safety, our key idea is to check each CPS component several times in succession. In the next section, using the formula for total probability and the Bayesian approach, we build a mathematical model for finding the minimum number of necessary repeated tests required for each component. In Sections 3 and 4, we develop an appropriate test scheduling algorithm and examine its properties. Finally, in Section 5, we describe computational experiments to detect failures in a real educational CPS serving for hybrid teaching; we compare the proposed algorithm with several known failure detection algorithms and obtain encouraging practical results. Section 6 concludes the work.

2. A FAILURE DETECTION PROBLEM AND THE MATHEMATICAL MODEL

Let a multi-component ECPS be known to have failed. An exact location of the failed part in the CPS is not known precisely, only preliminary probabilities may be estimated by experts. It is necessary to determine what is the exact location of the failed part. To do this, a system of unreliable sensors is used that sequentially checks the ECPS components. The failed part, which we shall refer to as the *target*, is "hidden" in one of a finite set I of possible locations (components), with a priori known probability p_i (where $\Sigma_{i\in I}\, p_i = 1$), the total number of all the possible components being $m = |I|$. The detecting sensors are smart but not totally perfect, and this sensor system is used by a searching team to sequentially inspect each location, possibly more than once, in order to find the "hidden target".

Associated with each location i, along with the above probability p_i, are the following given data:

- a cost c_i and time t_i for inspecting/searching location i;
- an "overlook probability" α_i, i.e., the probability that if the target is in location i and when the sensor inspects the location i once, it does not discover the target; in this case we say that there happened sensor's overlooking, or a *false-negative outcome*;
- a "false alarm" probability β_i; this is the probability that the location i does not contain the hidden target but an imperfect sensor erroneously classifies the location as containing the target we are looking for; in this case we say that a *false-positive outcome* occurs.

Our problem is to find a sequential strategy (i.e., a sequence) $\pi = [\pi(1), \pi(2), \ldots]$, where $\pi(k)$ is a location (or a component) inspected at the kth step in π, such that the expected cost $V(\pi)$ of finding the object is minimal. A precise value of the $V(\pi)$ will be defined below. Note that, since the sensor is imperfect, before the object is successfully found any location may be searched, in the worst case, more than once.

Following Matula [12], we call a strategy *ultimately periodic* if $\pi(j + \theta) = \pi(j)$ for all $j > T$, where T denotes the length of the *initial transient phase* and θ the *length of the period*. In his seminal work of 1964, Matula [12] had investigated a special case of the above problem where the false-positive inspection outcomes do not occur, and only false-negative outcomes are studied. He proved the existence conditions for a periodic optimal search sequence with the overlook inspection errors only and found how the initial overlook probabilities and costs affect the minimal period and transient phase length of the periodic sequence. The present paper extends the Matula search model. It treats the case in which the imperfect sensor may give both the false-positive and false-negative indications. We obtain the existence conditions for a periodic, asymptotically optimal search sequence and find the possible lengths of the period and the transient phase.

To handle the false-positive indications of the sensor, we introduce and compute the conditional probability, denoted by $p_{[i,h_i]}$, of the event that the hidden target is indeed in a location i under condition that the sensor has repeatedly detected the target in the location i during h_i sequential inspections, where the integer parameter h_i is chosen in such a way that the latter probability $p_{[i,h_i]}$ will be no less than the *success level SL*, a given parameter which is set in advance by the human searcher.

2.1. Basic Properties of the Search Process

We start with presenting the given input data in more detail (see Table 1).

Table 1. Given input data

Event C_i	{the target is hidden in a location i}
$P(C_i) = p_i$	a priori probability that the target is hidden in location i
Event B_i	{the sensor has classified a location (or, a component) i as a hidden target (either correctly, or probably not correctly) during a single inspection of the location i in sequence π}. In other words, the sensor testifies that it has detected a target in the location i as a result of a single inspection of this location at some step of the sequential strategy π, but we cannot be sure that this outcome is true since the sensor is imperfect
$P(C_i) = p_i$,	a priori probability that the target object is hidden in location i
$\alpha_i = P(\bar{B}_i \vert C_i)$	Probability that the target object is *not* detected in location i under condition that the object is actually hidden in location i); this is the *overlooking probability*, or a *false-negative outcome* mentioned in the Introduction
$\beta_i = P(B_i \vert \bar{C}_i)$	probability that a no-target object is erroneously classified by the sensor as the target when searching in location i although, in fact, the target object is *not* in location i; this is called the *false-alarm probability*, also referred to as the probability of a *false-positive inspection outcome*

Next, we introduce several auxiliary parameters depending on the above input data, which are presented in Table 2.

Table 2. Auxiliary parameters

$f_i = P(B_i)$	probability that the sensor has classified a location i as the hidden target during a single inspection of that location i at some step of the sequential strategy}
$P(B_i \vert C_i)$,	the conditional probability that the target is *correctly* detected by the sensor in location i during a single inspection of the i under condition that the hidden object is actually in location i
h_i	the integer whose role is explained in the following definitions
Event $\{C_i \vert B_i^{(1)} \cap B_i^{(2)} \ldots \cap B_i^{(h_i)}\}$	{the hidden target is really in location i under condition that during h_i sequential, but not necessarily consecutive, inspections of that location, the sensor has testified h_i times that it detected (correctly or maybe not correctly) that the target is in the location i, the integer parameter h_i being fixed}
$p_{[i,h_i]}$	$p_{[i,h_i]} = P\left(C_i \middle\vert B_i^{(1)} \cap B_i^{(2)} \ldots \cap B_i^{(h_i)}\right)$, the conditional probability of the above event $\{C_i \vert B_i^{(1)} \cap B_i^{(2)} \ldots \cap B_i^{(h_i)}\}$
SL	the required *safety level*; this is an additional parameter which is set by the searcher in advance as a permissible lower bound on the above probability $p_{[i,h_i]}$
H_i	the minimum amount (*min* h_i) of sequential inspections (not necessarily consecutive ones) of location i depending on the given input probabilities and a known SL value, and such that the $p_{[i,h_i]}$ value satisfies the above inequality (1): $H_i = \min (h_i) \vert p_{[i,h_i]} \geq SL)$. H_i is precisely computed below

Simply speaking, parameters h_i and SL are chosen in such a mode that the following inequality (1) holds:

$$p_{[i,h_i]} \geq SL. \tag{1}$$

Using the total probability formula and Bayes' formula, we can obtain that

$$f_i = P(B_i \vert C_i)P(C_i) + P(B_i \vert \bar{C}_i)P(\bar{C}_i) = (1 - \alpha_i)p_i + \beta_i(1 - p_i).$$

In order to precisely and efficiently compute the H_i values, we need the following claims:

Claim 2.1. The conditional probability, for a single search, is:

$$P(C_i|B_i) = \frac{P(B_i|C_i)P(C_i)}{P(B_i|C_i)P(C_i)+P(B_i|\bar{C}_i)P(\bar{C}_i)} = \frac{(1-\alpha_i)p_i}{(1-\alpha_i)p_i+\beta_i(1-p_i)}.$$

Claim 2.2. If an integer parameter h_i is given, the conditional probability $p_{[i,h_i]}$ can be found as follows:

$$p_{[i,h_i]} = P\left(C_i \,\middle|\, B_i^{(1)} \cap B_i^{(2)} \cap \ldots \cap B_i^{(h_i)}\right) = \frac{p_i(1-\alpha_i)^{h_i}}{p_i(1-\alpha_i)^{h_i} + (1-p_i)\beta_i^{h_i}}$$

Claim 2.3. For any location i, $i \in I$, the $H_i = \min_i h_i$ value can be found as follows:

$$H_i = \left\lceil \frac{log\left(\frac{p_i}{1-p_i} \cdot \frac{1-SL}{SL}\right)}{log\left(\frac{\beta_i}{1-\alpha_i}\right)} \right\rceil,$$

where $\lceil x \rceil$is the ceiling value of a real number x, that is the smallest integer number following x.

The proofs of Claims 2.1-2.2 are straightforward, and we omit them. The Claim 2.3 directly follows from the relation (1), Claim 2.2, and the following inequality:

$$p_{[i,h_i]} = P\left(C_i \,\middle|\, B_i^{(1)} \cap B_i^{(2)} \cap \ldots \cap B_i^{(H_i)}\right) = \frac{p_i(1-\alpha_i)^{H_i}}{p_i(1-\alpha_i)^{H_i}+(1-p_i)\beta_i^{H}} \geq SL.$$

The Stopping Rule. Our key assumption is that the searching process is allowed to stop when the bounding inequality (1) holds for some inspection, where the required value of *SL* is known, that is, the probability $p_{[i,h_i]}$ of successful detection of the target in some location *i* should be no less than the required success level *SL*. In other words, the search of the target finishes when the hidden target is detected by the sensor in some location *i* exactly H_i times, for some i, $i \in I$, the value H_i being computed in advance in Claim 2.3.

The relation (1), Claim 2.3, and the *Stopping Rule* display the principal difference between the Matula model [12] and the presented search process scenario. Namely, in the Matula model it is assumed that the hidden target is found immediately as soon as the sensor testifies (for the first time) that it has detected the hidden target during an inspection of some location, and immediately, after that this event occurs, the search process stops. This is correct because in this model the false-positive outcomes are absent.

In contrast, in the present scenario, a single detection of the target by the imperfect sensor is insufficient, and after the first detection of the target the search process should be continued; the search may be stopped only when, in a certain location, the target is discovered by the sensor 'sufficiently many times' (or, to say precisely, H_i times). We define the sufficient value H_i in Claim 2.3 in such a way that guarantees that the probability $p_{[i,h_i]}$ of a 'correct discovery' of the target reaches a required sufficiently high safety level *SL*, which is selected in advance by the searcher.

3. AN INDEX-BASED GREEDY ALGORITHM

Consider an initial sub-sequence of the strategies contains $H_1 - 1$ inspections of component 1, $H_2 - 1$ inspections of component 2, …, $H_m - 1$ inspections of component m, where all the 'heights' H_i are defined in Section 2. For instance, we can take the initial part of the strategies as follows:

$$U_{[\pi,0]} = \left\{ \underbrace{1,1,\dots,1}_{H_1-1\ times},\ \underbrace{2,2,\dots,2}_{H_2-1\ times},\dots,\underbrace{m,m,\dots,m}_{H_m-1} \right\}. \tag{2}$$

Let us consider the following sequential search strategy $\pi = \{U_{[\pi,0]}, \pi[1],\dots, \pi[n], \dots\}$. Such a choice of the initial sub-sequence $U_{[\pi,0]}$ is motivated by the stopping rule formulated above; namely, the $U_{[\pi,0]}$ provides that during the first $|U_{[\pi,0]}|$ steps of any strategy π, the search process will not stop and should be continued according to the greedy algorithm described in this section below.

We need the following additional notation:

- $M(i, N, \pi)$, the number of inspections of a location i among the first N steps of a sequence π under condition that (i) the sensor discovers the target in the location *i exactly H_i times* during the first N steps of sequence π, and (ii) the sensor discovers the target in that location i at the 'final' Nth step, after which the search stops;
- $\pi(N) = i$;
- $P_i = P(i, N, \pi) = P\big(M(i, N, \pi)\big)$, the probability that the sensor discovers the target in location *i exactly H_i times* among first N steps of sequence π, under condition that the target is discovered at the Nth step, after which event the search stops;
- t_i , time spent for inspecting the location i;
- c_i , cost assigned to a single inspection of a location i in the linear minimum-cost search model;
- $V(\pi)$, the total search cost assigned to strategy π. $V(\pi)$ is a random function depending on the random number of steps in strategy π before the search process stops; this function is defined in detail below.
- the search process in the sequence π stops at the μth step of sequence π, where μ is a random number that obtains the integer values 1,2,3,…

If $\mu = N$, this means that within the first N steps the sensor detects the target in location $\pi[N] = i$ exactly H_i times. Now we can formulate more precisely in which way this search process differs from the search scenario of Matula [12]. Actually, the search under investigation can be looked at as a sequence of independent Bernoulli trials where, in terminology of the Bernoulli trials, a 'success' corresponds to the event that the sensor claims that it detects the target during a single inspection. Denote by $P(\mu = N)$ the probability of H_i successes in location $\pi[N] = i$ occurring during $M(i, N, \pi)$ trials. It is known that the random variable μ has the negative binomial distribution (NBD) with the success probability f_i in a single trial. (For the detailed NBD definition and notation, we refer the reader to the standard texts, e.g., De Groot [5]) or Wentzel [16]. Therefore,

$$P(\mu = N) = P\big(M(i, N, \pi)\big) = \binom{M(i, N-1, \pi)}{H_i - 1} (1 - f_i)^{M(i,N-1,\pi)+1-H_i} \cdot f_i^{H_i}$$

Define the random cost $R(\mu, \pi)$ of the first μ steps of strategy π as follows:

$$R(\mu, \pi) = R(\pi[\mu], \mu, \pi) = c_{\pi[\mu]} \cdot T(\pi[\mu], \mu, \pi),$$

where $T(\pi[\mu], \mu, \pi)$ is the random search time before the search process stops, which is computed as follows: $T(\pi[\mu], \mu, \pi) = \sum_{j=1}^{m} t_j\big(H_j - 1\big) + \sum_{k=1}^{\mu} t_{\pi[k]}$. In this formula, the first term corresponds to the duration of the initial sub-sequence of steps $U_{\pi[0]}$ in sequence π, whereas the second term is the duration of all subsequent steps of π. Notice that $T(\pi[\mu], \mu, \pi)$ is random since μ is random.

Denote $T(\pi[N], N, \pi) = T(i, N, \pi)$ by T_i. Then the expected cost of the target search in the sequence π is:

$$V(\pi) = Exp\big(R(\mu,\pi)\big) = Exp\big(R(\pi[\mu],\mu,\pi)\big) = \sum_{N=1}^{\infty}\big(R(\pi[N],N,\pi)\big)\,P(\mu = N) =$$
$$= \sum_{N=1}^{\infty}\big(R(i,N,\pi)\big)\cdot P(\mu = N) = \sum_{N=1}^{\infty} c_i \cdot T_i \cdot P(\mu = N) =$$
$$= \sum_{N=1}^{\infty} c_i \cdot T_i \cdot \binom{M(i,N-1,\pi)}{H_i - 1}(1-f_i)^{M(i,N-1,\pi)+1-H_i} \cdot f_i^{H_i}$$

Let us define an '*attractiveness*' of each location, which is inspected after the initial sub-sequence $U_{[\pi,0]}$, as the following preference ratios Q_i:

$$Q_i = \frac{c_i \cdot P_i}{t_i} = \frac{c_i \cdot P_i\big(M(i,N,\pi)\big)}{t_i} = \frac{c_i \cdot \binom{M(i,N-1,\pi)}{H_i-1}(1-f_i)^{M(i,N-1,\pi)+1-H_i} \cdot f_i^{H_i}}{t_i}. \tag{3}$$

The following claim permits to define the optimal (i.e., minimum-cost) search sequence.

Theorem 3.1. The strategy π is optimal iff at each step N after the initial sub-sequence $U_{[\pi,0]}$ the next inspected location $\pi(N) = i*$ is selected in such a way that its ratio $Q_{i*} = \max_j Q_j$, where the ratios Q_j are defined in (3).

Notice that, when selecting a next location $\pi(N) = i^*$, the maximum ratio Q_{i*} is calculated for all $j \in I$ using the 'new' values of $P_i = P_i\big(M(i,N,\pi)\big)$ calculated at each new step. The proof of Theorem 1 is straightforward by the standard 'interchange argument' well known in scheduling theory (see, e,g, Blazewicz et al. [4], Lenstra and Shmoys [10]) and is left to the reader. This claim extends Theorem 1 of Matula [12], which is valid when all the false-positive probabilities are zeros; a minor technical difference is that in place of the times t_i used in the present model, Matula's model uses costs c_i, for all i. The attractiveness ratio Q_i depends on the outcomes $M(i,N,\pi)$ of already-done search steps – which is not typical for standard scheduling problems without task repetitions, wherein the ratios Q_i do not depend upon a 'history' of the search.

In what follows, we will be interested in investigating the periodicity property of the optimal solutions. For this aim, we need to introduce the definitions of asymptotically suitable and asymptotically optimal solutions. Consider a sequence π and the first N searches in the π. Let i denote the location inspected at the Nth step of π: $\pi(N) = i$. We shall call an inspection sequence *suitable* if all locations in it are ordered as pointed out in Theorem 3.1. The sequence π is called *asymptotically suitable* if all the locations in it are ordered in such a way that, at any step N of π made after the initial sub-sequence $U_{[\pi,0]}$ the next inspected location $\pi(N) = i$ is selected in such a way that its preference ratio Q_i satisfies the following: $Q_i/\max_j Q_j = 1-\alpha(N)$, where the max-operator is taken over all the preference ratios computed at the step N, and $\alpha(N) \to 0$ when $N \to \infty$.

A sequence π^* is called *asymptotically optimal* if, for this sequence π^* and the optimal sequence s^{opt}, the following holds:

$$\lim_{N\to\infty}\Big(\Delta V(\pi^*, N) - \Delta V\big(s^{opt}, N\big)\Big) = 0, \tag{4}$$

where $\Delta V(s, N)$ denotes the value of the cost function computed for the "tail" of the infinite sequence s, where s stands for π^* or s^{opt}, beginning with the $(N+1)$th step. This definition means that the difference between the values of the 'tails' of sequences $\pi^*$and s^{opt} is arbitrarily small for sufficiently large N and approaches zero when $N \to \infty$.

We assume that there exists the 'learning effect' during the search. It means that, if an inspection is repeated several times in the same location, the inspection duration decreases for larger steps. Specifically, in the claim below the inspection time decreases faster than $1/N^2$ with the growth of N.

Theorem 3.1'. If (i) the durations of inspections are decreasing for larger steps as $t_i(N) = O(1/N^{2+\varepsilon})$, for any $\varepsilon > 0$ and for each $i \in I$, and (ii) π^* is asymptotically suitable, then $\pi^*$is asymptotically optimal.

The claim immediately follows from the definitions of asymptotically suitable and asymptotically optimal sequences and the description of the learning effect.

4. PERIODICITY PROPERTIES OF THE GREEDY SEARCH STRATEGY

Let us determine under which conditions an ultimately periodic, asymptotically optimal sequence exists for the failure detection problem with false-positive and false-negative outcomes. The next Lemma and Corollary are analogous to the similar claims in Matula [12]. The proof of these facts is along the same line as that in [12].

Lemma. If π is an ultimately periodic, asymptotically optimal sequence of transient length T and period $\theta = \sum_{i \in I} \sigma_i$, where σ_i is the number of inspections of location i per period, then $(1 - f_i)^{\sigma_i} = (1 - f_j)^{\sigma_j}$, for any $i,j \in I$.

Corollary. A necessary condition for the existence of an ultimately periodic and asymptotically optimal sequence in the case of two error types is that the set of ratios $\left\{ \frac{log(1-f_i)}{log(1-f_j)} \right\}, i, j \in I$ consists only of rational numbers.

The next theorem highlights the impact of two error types on the parameters of ultimately periodic, asymptotically optimal solutions.

Theorem 4.1. For the considered search problem with two error types, where the ratios $\left\{ \frac{log(1-f_i)}{log(1-f_j)} \right\}$ are rational numbers for $i, j \in I$, there exists a search sequence π^* such that:

(a) π^* is ultimately periodic of period θ and *initial transient length* denoted by T,

$$\theta = min \left\{ \theta' \middle| \theta' \text{ and } \frac{\theta'}{\sum_{j \in I} \left(\frac{log(1-f_i)}{log(1-f_j)} \right)} \text{ are integers for } i \in I \right\}, \tag{5}$$

$$T = \sum_{i \in I} \left[\min_{n = H_i - 1, H_i, \ldots} \left\{ n \middle| 0 < \frac{c_i \cdot \binom{n}{H_i - 1} (1 - f_i)^{n+1-H_i} f_i^{H_i}}{t_i} \le \min_{j \in I} \left\{ \frac{c_j f_j^{H_j}}{(1 - f_j)^{H_j} t_j} \right\} \right\} \right], \tag{6}$$

(b) If the searching process involves the 'learning effect' described in Theorem 3.1', the sequence π^* is asymptotically optimal,

(c) θ is the minimal possible period (in the periodic part of the sequence π^*).

Here T denotes an *initial transient length*, which will be upgraded to $\tilde{T}$ in Theorem 4.1*(a) below.

Remark 4.1. Meaning of formula (5) is that the value of period θ and the numbers of inspections of each location within the θ should be integer and will guarantee the minimum to the θ. The structure of the obtained solution and the stopping rule are notably different in comparison with the single-type error model by Matula [12].

For the search problem under investigation there exists at least one suitable sequence π, satisfying (3) and Theorem 3.1. Let us introduce the following notation:

$$\phi(N) = \max_{i\in I}\left\{\frac{c_i \cdot \binom{M(i,N-1,\pi)}{H_i-1}(1-f_i)^{M(i,N-1,\pi)+1-H_i} \cdot f_i^{H_i}}{t_i}\right\}, \qquad N = 1,2,... \tag{7}$$

$$Y = \min_{j\in I}\left\{\frac{c_j \cdot f_j^{H_j}}{(1-f_j)^{H_j} t_j}\right\}, \tag{8}$$

$$T = \min\{N|\phi(N) \le Y\}, \tag{9}$$

$$\sigma_i = \frac{\theta}{\sum_{j\in I}\frac{log(1-f_i)}{log(1-f_j)}}, i \in I, \tag{10}$$

$$X = \exp\left\{\frac{\theta}{\sum_{j\in I}\frac{1}{log(1-f_j)}}\right\}, \tag{11}$$

$$\text{and } K = min\{K'|\phi(K') \le XY\}. \tag{12}$$

The definitions of T in (6) and (9) are equivalent (see the proof of Claim 4.1 in Section 5). Using (10) and (11), we obtain:

$$X = exp\left\{\frac{\theta}{\sum_{j\in I}\frac{1}{log(1-f_j)}}\right\} = exp\left\{\frac{\sigma_i \sum_{j\in I}\frac{log(1-f_i)}{log(1-f_j)}}{\sum_{j\in I}\frac{1}{log(1-f_j)}}\right\} = exp\{\sigma_i\, log(1-f_i)\} = (1-f_i)^{\sigma_i}. \tag{13}$$

Since the numbers $\left\{\frac{log(1-f_i)}{log(1-f_j)}\right\}$ are rational, $\sum_{j\in I}\frac{log(1-f_i)}{log(1-f_j)}$ is also rational for any $i \in I$, giving meaning to the relation (5).

It is easy to see that, for any i, the ratio $Q_i\big(M(i,N-1,\pi)\big)$ defined in (3) is monotonically non-increasing function of $M(i,N-1,\pi)$ that approaches zero. The latter fact gives meaning to parameter R and inequality (14) defined below.

Let us show that in the considered search scenario, along with the transient phase, which is analogous to Matula's transient phase, the asymptotically optimal strategy has a so-called *preamble phase* that precedes the transient phase and is defined as follows. Denote by R_i the minimum value of $M(i,N-1,\pi)$ for which the function $Q_i\big(M(i,N-1,\pi)\big)$ is a monotonically non-increasing and, in addition, the following inequality holds:

$$\binom{R_i}{H_i-1}(1-f_i)^{R_i+1} < 1 \tag{14}$$

Consider R_i consecutive initial searches of any location i, which form a part of a *preamble phase of the location i.* Consider the concatenation C^m of m individual preliminary searches for all the locations, each one containing R_i consecutive searches of location i, i=1,2,…, m. Totally, this initial part of π has $R = \sum_i R_i$ inspections. Finally, define the *preamble* sub-sequence S^{pr} of π to be a longer subsequence among sub-sequences, C^m and $U_{s(0)}$, which contains $\widehat{R}$ = max (R, | $U_{s(0)}$) |) steps, where $U_{s(0)}$ was defined in (2) in Section 3.

From the definition of S^{pr}, it follows that after that the preamble phase S^{pr} of any search sequence is done, the following properties hold:

1. The function $\phi(N)$ is a monotonically non-increasing function that approaches zero.
2. $\binom{M(i, N-1, \pi)}{H_i - 1}(1-f_i)^{M(i,N-1,\pi)+1} < 1$, for any location i and for each step N after the preamble phase.
3. If the false positive probabilities are zeros (i.e., $H_i = 1$ for each location i) then R = 0.

Now let us define the set

$$G = \left\{ i \,\middle|\, \frac{c_i \cdot f_i^{H_i}}{(1-f_i)^{H_i} \cdot t_i} = \Upsilon \right\} \qquad (15)$$

and let us arrange all the elements of G as follows: $G = \{i_0, i_1, i_2, \ldots . i_L\}$.

Now we are in position to define the desired sequence π^*. Let π be some suitable sequence. Define a sequence π^*as follows:

$\pi^* = \{S^{pr}\}$, for all locations from step 1 up to step $\widehat{R}$ = max (R, | $U_{s(0)}$) |);
$\pi^*(j) = \pi(j)$ for $\hat{R} + 1 \le j \le \hat{R} + K - 1$;
$\pi^*(j + \hat{R} + K) = i_j$ for $0 \le j \le L, i_J \in G$;
$\pi^*(j + \theta) = \pi^*(j)$ for $j > \hat{R} + K + L - \theta$.

From the above construction of the sequence π^*, the transient phase in π^* terminates when the first period θ in the periodic part of π^* starts; hence, it has the length $\hat{T} = \hat{R} + K + L + 1 - \theta$.

We can now prove the following:
Theorem 4.1' (a). π^* is ultimately periodic with period θ and *augmented transient length* $\tilde{T}$, where

$$\hat{T} = \hat{R} + K + L + 1 - \theta = \hat{R} + T - \sum_{i \in G}\left(\left\lceil log_{(1-f_i)}\binom{M(i, K-1, \pi^*)}{H_i - 1}\right\rceil\right) \qquad (16)$$

Notice that, in comparison with the parameter $\tilde{T}$ in (16), the transient length T in the Matula work does not contain the preamble term $\hat{R}$ nor the negative term with the sum of logarithms.

Theorem 4.1(b). If the searching process involves the 'learning effect', π^* is asymptotically optimal.

We prove by induction on k (where k=0,1,2,…) that π^* is suitable (and, hence, asymptotically suitable) through stages from $\hat{R} + 1$ to $\hat{R}$ + K + L + k.

For k=0 the required claim follows from the definition of the sequence $\pi^*$and the fact that the sequence π is suitable. We have:

$$\pi^*(j) = \pi(j), \text{if } \hat{R} \le j \le \hat{R} + K - 1.$$

Next, by virtue of (16) we have that

$$\phi(K) = XY = \frac{c_i \cdot \binom{M(i, K-1, \pi^*)}{H_i - 1} (1-f_i)^{M(i,K-1,\pi^*)-H_i+1} \cdot f_i^{H_i}}{t_i}, \qquad \text{if } i \in G,$$

$$\text{and, for any } j,\ \frac{c_j \cdot \binom{M(j,K-1,\pi^*)}{h_j-1} (1-f_j)^{M(j,K-1,\pi^*)+1-H_j} \cdot f_j^{H_j}}{t_j} = \phi(K) \le XY.$$

Therefore, the locations from the set G are, indeed, the locations of the *suitable* sequence π^*:

$\pi^*(K) = \pi(K) = i_0, \pi^*(K+1) = \pi(K+1) = i_1, \pi^*(K+2) = \pi(K+2) = i_2, \ldots, \pi^*(K+L) = \pi(K+L) = i_L.$

By virtue of Theorem 3.1', the sequence π^* is asymptotically optimal up to stage $K+L$.

Assume now that $\pi^*$is suitable through stage $K+L+k$ and let $\pi^*(K+L+k+1) = i$, hence also $\pi^*(K+L+k+1-\theta) = i$ (for the simplicity of notation, here we omit symbol $\hat{R}$).
Then

$$\frac{c_i \cdot \binom{M\left(i, K+L+k, \pi^*\right)}{H_i - 1} \left(1-f_i\right)^{M\left(i,K+L+k,\pi^*\right)+1-H_i} \cdot f_i^{H_i}}{t_i} =$$

$$= \frac{\left(1-f_i\right)^{\sigma_i} \binom{M\left(i, K+L+k, \pi^*\right)}{H_i - 1}}{\binom{M\left(i, K+L+k, \pi^*\right) - \sigma_i}{H_i - 1}} \cdot \frac{c_i \cdot \binom{M\left(i, K+L+k, \pi^*\right) - \sigma_i}{H_i - 1} \left(1-f_i\right)^{M\left(i,K+L+k,\pi^*\right)-\sigma_i+1-H_i} \cdot f_i^{H_i}}{t_i} =$$

$$= \frac{\binom{M\left(i, K+L+k, \pi^*\right)}{H_i - 1}}{\binom{M\left(i, K+L+k, \pi^*\right) - \sigma_i}{H_i - 1}} \cdot \max_j \frac{\binom{M\left(i, K+L+k, \pi^*\right) - \sigma_i}{H_i - 1}}{\binom{M\left(i, K+L+k, \pi^*\right)}{H_i - 1}} \max_j \frac{c_j \cdot \binom{M\left(i, K+L+k, \pi^*\right)}{H_i - 1} \left(1-f_j\right)^{M\left(j,K+L+k,\pi^*\right)+1-H_j} \cdot f_j^{H_j}}{t_j} =$$

$$= \left(1+\alpha_1\left(M\left(j, K+L+k, \pi^*\right)\right)\right)\left(1+\alpha_2\left(M\left(j, K+L+k, \pi^*\right)\right)\right) c_j \cdot \binom{M\left(i, K+L+k, \pi^*\right)}{H_i - 1} \left(1-f_j\right)^{M\left(j,K+L+k,\pi^*\right)+1-H_j} \cdot f_j^{H_j} =$$

$$= \max_j \frac{c_j \cdot \binom{M\left(j, K+L+k, \pi^*\right)}{H_j - 1} \left(1-f_j\right)^{M\left(j,K+L+k,\pi^*\right)+1-H_j} \cdot f_j^{H_j}}{t_j} \left(1+\alpha\left(M\left(j, K+L+k, \pi^*\right)\right)\right)$$

The latter relations show that $\pi^*$is asymptotically suitable, and, hence, asymptotically optimal through stage $K+L+k+1$, and by induction axiom asymptotically optimal at all stages.

By virtue of Theorem 4.1 and relation (16), the period length θ and the augmented transient length $\tilde{T}$ do not exceed $3m$ (recall that m is the number of components) Therefore, if the number of periods $\mathfrak{C}$ is fixed, then the number of steps of the greedy algorithm does not exceed O($\mathfrak{C}m$), and the total computing time does not exceed O($\mathfrak{C}m$).

5. NUMERICAL EXPERIMENTS AND COMPARISONS

In this section, we briefly report about our computational experiments for detecting failures in a real educational CPS serving for hybrid teaching. We compare the results produced by the proposed algorithm with several known failure detection algorithms and obtain encouraging practical results. As an illustration, consider a sample experiment in which we have selected three most vulnerable parts of the ECPS; the input data are presented in Table 1; SL = 0.95 as selected by a decision maker.

Table 1. Input data

	Part 1	Part 2	Part 3
$p_i = P(C_i)$	0.08	0.28	0.64
$\beta_i = P(B_i \mid \bar{C}_i)$	0.46	0.67	0.57
$\alpha_i = P(\bar{B}_i \mid C_i)$	0.05	0.05	0.15
t_i (Min.)	5	8	10
c_i ($)	100	10	1

Applying the formulas for probabilities f_i, heights H_i, and ratios Q_i, we find that f_1=0.50; f_2=0.75; f_3=0.75; H_1=2, H_2 = H_3=1, and U_{s0} =<1>, $\hat{T} = \theta$ = 4. Omitting elementary computations, we have that the transient sequence is (1,1,1,2), which is followed periodically by the sequence (3,2,1,1), that is, we have found that the best search sequence, up to the 8th step, is as follows: π^{opt} =<1,1,1,2,(3,2,1,1)>, its objective function value being $F(\pi^{opt})$ = 2,157.78. If we take, for comparison, a random search sequence, for instance, s^{random} = <1,1,1,2,(2,3,2,1)>, we obtain a worse value: $F(s^{random})$=2, 257.78. If we continue the calculations, we obtain that, at Step 24, $F(\pi^{opt})$ = 2,657.68 and $F(s^{random})$=3,012.47. After this step, further changes in the objective values become negligible.

It should be noted that the proposed fast algorithm, which guarantees that the probability of correctly detecting a target is no lower than a given safety level, in practice outperforms several known combinatorial failure detection algorithms. Namely, in our simulations, we have consistently observed that the Matula algorithm stops as soon as the “first success” occurs, that is, when a sensor first detects that a part is faulty. Such a ‘solution’ is not applicable to the scenario considered in this paper with possible false alarms. Another closely related algorithm is Alidaee’s algorithm [1], in which sensors are allowed to repeat the inspections of vulnerable components, and the search process continues until either sensor declares twice that a certain part is faulty. For the convenience of the reader, we present the following formula for the probability ofp_i a successful target detection by the Alidaee algorithm:

$$\frac{p_i(1\text{-}\alpha_i)^2}{p_i(1\text{-}\alpha_i)^2 + (1\text{-}p_i)_{\beta_i^{\ 2}}} .$$

In the considered example, this probability is smaller than the given level SL; in general, it may be notably smaller than the probability of correct detection provided by the proposed algorithm.

6. CONCLUSIONS

Various probabilistic and statistical methods are widely known today for assessing the probability of failure detection in large-size systems, depending on the reliability of imperfect sensors. Our approach is based on probabilistic safety assessment.

The developed framework aims to be a part of the Education 4.0 concept, where the authors utilize Industry 4.0 technologies to efficiently and quickly detect the faulty parts in the CPSs, by this improving the efficiency of educational process and increasing the perception of the studied material [3, 13].

Future work on the proposed approach will be to further integrate Industry 4.0 technologies with the 'teaching factory'. In addition, promising directions for future research are further development and extension of the proposed fast failure-detection algorithm for more general scenarios such as multiple failures, multiple parallel searchers, as well as prevention and mitigation of possible risks caused by failures. Algorithm comparison will be continued and expanded in our future research.

Acknowledgement

The second author gratefully acknowledges the support by a grant from Ariel University and Holon Institute of Technology (Israel) "Implementation of artificial intelligence methods to improve early detection of disease outbreaks, prevention and management" (no. RA 220-10).

References

[1] B. Alidaee. "*Optimal ordering policy of a sequencing model*", Journal of Optimization Theory and Applications, vol. 83(1), pp. 199-205, (1994).
[2] S. Alpern, R. Fokkink, L. Gasieniec, R. Lindelauf, and V.S. Subrahmanian (Eds.). "*Search Theory",* Springer, 2013, Berlin.
[3] S. Bachir and A. Abénia. *Internet of everything and educational cyber physical systems for University 4.0*, Lecture Notes in Computer Science, vol. 11684 (2019).
[4] J. Blazewicz, K.H. Ecker, E. Pesch, G. Schmidt, M. Sterna, and J. Weglarz. *Handbook on Scheduling, From Theory to Practice,* Springer Nature, 2020, Berlin.
[5] M.H. De Groot, "*Optimal statistical decisions*". John Wiley & Sons, 1986, New York.
[6] J.B. Kadane. "*Optimal discrete search with technological choice*", Mathematical Methods of Operations Research, vol. 81, pp. 317–336 (2015).
[7] M. Kress, K.Y. Lin, and R. Szechtman. "*Optimal discrete search with imperfect specificity*", Mathematical Methods of Operations Research, vol. 68(3), pp. 539-549. (2008).
[8] B. Kriheli and E. Levner. (2013) "*Search and detection of failed components in repairable complex systems under imperfect inspections*", Lecture Notes in Computer Science, vol. 7630, pp 399–410, (2013).
[9] B. Kriheli, E. Levner, and A. Spivak. "*Optimal search for hidden targets by unmanned aerial vehicles under imperfect inspections*", American Journal of Operations Research, vol.6, pp. 153-166, (2016).
[10] J.K. Lenstra and D. Shmoys. *Elements of Scheduling*, Centrum Wiskunde & Informatica, 2020, Amsterdam.
[11] E. Levner. "*Infinite-horizon scheduling algorithms for optimal search for hidden objects*", International Transactions in Operational Research, vol. 1, pp. 241-250, (1994).
[12] D. Matula. "*A periodic optimal search*", The American Mathematical Monthly, vol. 71(1), pp. 15-21, (1964).
[13] D. Mourtzis, E. Vlachou, G. Dimitrakopoulos, and V. Zogopoulos. "*Cyber-physical systems and education* 4.0*–the teaching factory* 4.0 *concept"*. Procedia Manufacturing, vol. 23, pp.129-134, (2018).
[14] S. Roy and S.K. Das (Eds,). "*Principles of cyber-physical systems. An interdisciplinary approach*", Cambridge University Press, Cambridge, 2020.
[15] A. Sotoudeh-Anvari, S. Sadjadi, S. Molana, and S. Sadi-Nezhad. "*A new MCDM-based approach using BWM and SAW for optimal search model*", Decision Science Letters , vol. 7(4), pp. 395-404, (2018).
[16] E.S. Wentzel. (1982*). "Probability Theor*y", Mir Publishers, 1982, Moscow.