future internet

MDPI

Article

A Bayesian Approach Based on Bayes Minimum Risk Decision for Reliability Assessment of Web Service Composition

Yang Song, Yawen Wang * and Dahai Jin

State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China; song_yang@bupt.edu.cn (Y.S.); jindh@bupt.edu.cn (D.J.)
* Correspondence: wangyawen@bupt.edu.cn

Received: 9 November 2020; Accepted: 3 December 2020; Published: 4 December 2020

check for updates

Abstract: Web service composition is the process of combining and reusing existing web services to create new business processes to satisfy specific user requirements. Reliability plays an important role in ensuring the quality of web service composition. However, owing to the flexibility and complexity of such architecture, sufficient estimation of reliability is difficult. In this paper, the authors propose a method to estimate the reliability of web service compositions based on Bayes reliability assessment by considering it to be a decision-making problem. This improves the testing efficiency and accuracy of such methods. To this end, the authors focus on fully utilizing prior information of web services to increase the accuracy of prior distributions, and construct a Markov model in terms of the reliabilities of the web composition and each web service to integrate the limited test data. The authors further propose a method of minimum risk (MMR) to calculate the initial values of hyperparameters satisfying the constraint of minimal risk of the wrong decision. Experiments demonstrate that the proposed method is capable of efficiently utilizing prior module-level failure information, comparing with the Bayesian Monte Carlo method (BMCM) and expert scoring method (ESM), when the number of failures increased from 0 to 5, reducing the required number of test cases from 19.8% to 28.9% and 6.1% to 14.1% separately, improving the reliability assessment of web service compositions, and reducing the expenses incurred by system-level reliability testing and demonstration.

Keywords: web service composition; hyperparameter; minimum risk; reliability assessment; bayes reliability assessment

1. Introduction

Web services are service-oriented architecture technologies that are executed using standard web protocols to ensure their operation varying platforms. With the continuous maturity of service-oriented technology and the promotion of service-oriented architecture (SOA) [1], multiple atomic web services have been implemented to achieve more powerful service compositions while satisfying relevant rules and addressing the demands of businesses. To maintain the quality of web service composition and utilize testing resources reasonably, assessment methods for web service compositions are required to exhibit high error detection capabilities, low cost consumption, and broad applicability [2].

In software reliability testing, the purpose of reliability growth testing is to identify and incorporate corrective actions that improve the reliability of the system. Following the completion of growth testing, a reliability demonstration is performed to verify a specific reliability requirement. Both broad and deep studies have been conducted on software reliability assessment to expose software defects and improve software reliability at early stages. When the number of test cases is very high or the test data are limited, software input is selected using statistical methods to promote rapid reliability growth of the software to ensure high reliability software demonstration testing [3]. In this context, Bayes reliability inference has been extensively applied to reliability estimation, and a wealth of experts and prior knowledge about system/subsystem performances are associated with it [1]. The Bayesian approach has been adopted because it combines subjective judgment or prior experience with data obtained from test samples. In other words, it uses a combination of existing experience and new test data to assess reliability metrics. In this approach, a posterior distribution is derived based on a prior distribution and a likelihood function, and reliability testing is conducted based on this derived distribution. Following the addition of new sample data, this posterior distribution is employed as a prior distribution during the process of constructing a new posterior distribution. Under the Bayesian framework, historical data are usually used for the analysis of prior distributions with certain parameters [4]. However, the origin of the required prior distributions remains a major concern in Bayesian reliability analysis [5]. The appropriate choice of priors is critical to capture useful information present in historical data. The fundamental advantage of adopting the Bayesian approach is that it augments the quality of data, thereby reducing the uncertainty involved in decision making.

Thus, in this case, the prior information is derived from historic data or experts' suggestions. The two types of information regarding unknown parameters can be expressed in the form of statistical distributions, from which further statistical inferences can be made. Then, the posterior distribution is evaluated by integrating the prior distribution and the likelihood function constructed based on the sampling distribution of the data. Bayesian estimation with prior distribution has been extensively studied in the literature, for the reliability assessment, some studies are shown in Table 1. The integration of multiple prior conditions, especially when these prior conditions are derived from various unique information sources, is challenging in Bayesian inference. The challenge is to integrate these distributions in a reasonable and effective manner and obtain a single probability distribution by combining all prior information and knowledge. Anderson-Cook [6] highlights the problem that "If the authors have multiple small data sets that are each individually insufficient to answer the question of interest, then combining them and incorporating engineering or scientific understanding of the process should allow us to extract more from that collection of data." However, the models need to be carefully considered and evaluated to ensure that they accurately reflect the data and the underlying physical processes. The models have to be simple enough that they can be distinguished by the data, but at the same time complex enough to capture the physical processes. Jacobs et al. [7] studied an averaging method that combines the opinions of different experts. However, it does not adequately explain the deviations of different opinions and require a more complex model such as Bayesian model. In [8], the Bayesian approach is introduced to reliability using several examples, but not suitable for the web service composition. Considering both system and subsystem level data, other methods were proposed, such as the average multiple priors [9], pooling method [10], and integrating the derived prior through the system structure [11]. In [12] an integration-based method is proposed to estimate and pool the weighted priors; it utilizes Bayes' theorem to integrate heterogeneous priors. In addition, in the service composition concept, where the nodes are traversed in a given order to provide a service, some related techniques have been applied to solve the raised reliability issues, (e.g., Markov chain [13], Stochastic Reward Networks [14], Game theory [15], and Universal Generating Functions [16]). In the case of a complex system comprising a large number of subsystems and components in the field of reliability engineering, prior knowledge of system reliability can provide information about the subsystem and system levels. In the case of web service composition, system-level module reliability is dependent on subsystem module reliability performance. Therefore, utilizing the full range of existing reliability prior information and testing data at subsystem levels and integrating such multilevel prior information during testing is a challenging task.

The calculation of hyperparameters in priors is another aspect that needs to be addressed. If there are no reasonable grounds to specify the hyperparameters of the priors, the advantage of Bayesian analysis can be compromised due to the introduction of additional uncertainty. Unfortunately, the assumption of arbitrary values for the hyperparameters of Bayesian priors is not uncommon [17].

Other existing studies on hyperparameters, such as the invariance principle [18], maximum entropy [19], and Jeffrey principle [20], concern pure heuristic methods, which require large amounts of data and are difficult to confirm, thereby making them unsuitable for practical application [21]. The evaluation should be based on the complete range of prior information, and ignoring this issue may result in costly and inaccurate reliability testing.

Table 1. Different methods on the reliability assessment.

Authors	Year	Approaches
Anderson-Cook [6]	2009	By considering the totality of data simultaneously, instead of performing analyses on each data type separately, for decision-making; however, it does not apply the service composition.
Csenki, A [7]	2009	The authors studied an averaging method that combines the opinions of different experts. However, it does not adequately explain the deviations of different opinions and require a more complex model such as the Bayesian model
Moala F A, Rodrigues J, Tomazella V L D [8]	2009	The authors compare the posterior densities of the reliability function by using several examples.
M. Burgman, M. McBride, R. Ashton [9]	2011	The study proposed a method of multiple prior information integration is to average all multiple priors. However, the deviations of the different opinions are not properly quantified in the averaging approach.
A. OHgan, C. Buck, A. Daneshkhah [10]	2006	Linear and geometric pooling methods allow unequal weights for prior information and focus on multiple priors; however, it over depends the experts' judgement.
K. McConway [11]	1978	The study considered system configuration and structure; the system prior can be derived from subsystem priors.
Li Z S, Guo J, Xiao N C, et al. [12]	2017	The study proposed an integration-based method to estimate and pool the weighted priors; it utilizes Bayes' theorem to integrate heterogeneous priors.
Pham, N. H. Tran, S. Ren, W. Saad and C. S. Hong [13]	2020	A sampling-based Markov approximation (MA) approach is proposed to solve the combinatorial NP-hard problem, to overcome this issue of requires a long convergence time.
Di Mauro, M. Longo, F. Postiglione, and M. Tambasco [14]	2017	For a chain of network nodes in Service Function Chains, a double-layer model is adopted, where Reliability Block Diagram describes the high-level dependencies among the architecture components, availability analysis is carried out to characterize the minimal configuration of the overall system.
Bian, X. Huang, Z. Shao, X. Gao and Y. Yang [15]	2015	The authors proposed the DISCCA algorithm that guides the service nodes towards the Nash Equilibrium with short latency and low congestion, through decision making by individual users with local information to improve the reliability of system.
M. Di Mauro, M. Longo, and F. Postiglione [16]	2018	For Service Function Chaining, propose a Universal Generating Function (UGF) approach, which minimizes deployment cost while respecting a given availability requirement.

In this paper, the authors propose a reliability assessment method for web service compositions based on Bayes theory by fully utilizing prior information of web services. The authors propose the method considering two stages of web service compositions—reliability growth stage and reliability demonstration stage—building the reliability test frame for web service composition. Based on the reliabilities of various web services, the authors constructed a Markov model to calculate the reliability of the target web service composition.

The model both reflects the different in reliability growth and shows the feature of web service composition. To improve the efficiency of the estimation, depending on the failure information of web services to estimate the prior distribution of web service composition, the authors focused on solving the prior distribution and propose a method of minimum risk (MMR) to calculate hyperparameters and reduce the sample size adequately to meet a product's reliability specifications. Experiments demonstrate that the proposed method is capable of completely utilizing prior failure information during web service testing, decreasing the number of test cases, and improving testing efficiency.

The remainder of this paper is organized as follows. The reliability assessment model for web service compositions based on the Bayes method is presented in Section 2.1. The assessment method for web service compositions and the hyperparameter solving method based on MMR are discussed in Section 2.2. The experiments and their results are discussed in Section 3. Finally, the paper is concluded in Section 4.

2. Materials and Methods

2.1. Web Service Composition Reliability Model

Web service composition is the process of combining different services into a single service to perform more complex functions. Owing to the lack of prior system-level information and experimental data for web service compositions, the system reliability is required to be assessed based solely on subsystem-level information. In this section, an assessment method based on a system structure with subsystem-level priors and data is introduced. The authors utilized the existing reliability prior information and testing data at sub-web services level, and integrated them to evaluate the web service composition. The authors defined an abstract web service to be a set of web services equipped with the same function, and specific web services were selected from the abstract web service to integrate the composition. Figure 1 illustrates that the web service composition includes abstract web service candidates, W_i, where $W_i = \{w_{11}, w_{12}, \ldots, w_{1n}\}$ and each w denotes a specific web service, and that optimal web services are selected from a set of functionally equivalent web service candidates.

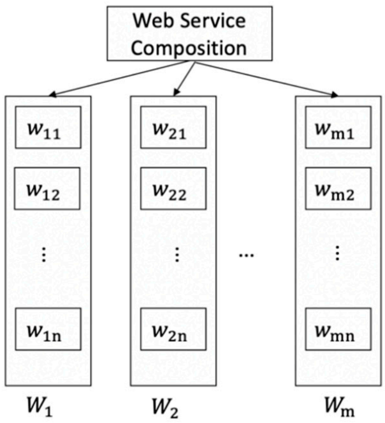

Figure 1. Web service composition.

The web service composition integrates specific web services following a certain logic. Assuming that all alternative web service candidates are available, the communication links for user-invoked web services depend on quality of service (QoS) properties including response-time, location, and latency, and using this, the invoked web service sets: $\{w_{11}, w_{22}, w_{32}, w_{43}\}$, illustrated in Figure 2, can be obtained.

To record the invocation of web services, the authors constructed a real-time data processing system, as depicted in Figure 3. The system includes Apache Kafka, Spark Streaming and database (DB), and web services deployed on Nginx. The invocation information of the web services was monitored and collected to calculate the associated transition probabilities. Among the constituent systems, Apache Spark exhibited high performance for both batch and streaming data, and Spark Streaming is an extension of the core Spark API that enables scalable, high-throughput, fault-tolerant stream processing of live data streams. The purpose of Kafka is to accumulate and, subsequently, re-compute the data. The authors utilized Spark Streaming to obtain data from the log of web services, which were then transmitted to the computation engine. Spark streaming was used as the computation engine, which directly programs based on SQL, to count the number of invocations and the number of errors corresponding to each web service. Considering the high dimensionality of queries, the results were stored on the OLAP database.

The authors used the aforementioned real-time data processing system to record the telemetry data regarding the number of invocations and the number of errors corresponding to each web service. In this paper, the authors considered the overtime of response as web service failure and used the proportion of invocation failures to calculate the reliability of web services. Let us assume that corresponding to a duration $t = t_n - t_k$, the numbers of invocations and failures of the web service w is C_n and n, respectively, at time t_n, and C_k, and k, respectively, at time t_k. Therefore, the probability of failure of the web service w corresponding to the duration t is $F(w) = \frac{n-k}{C_n - C_k}$, and its reliability is $r(w) = 1 - F(w)$.

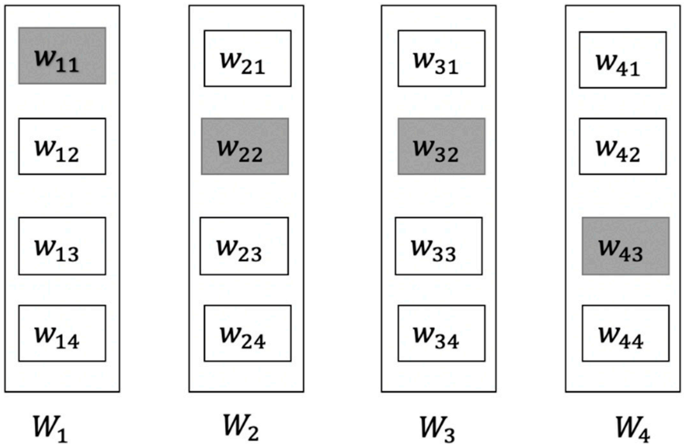

Figure 2. Web service flow.

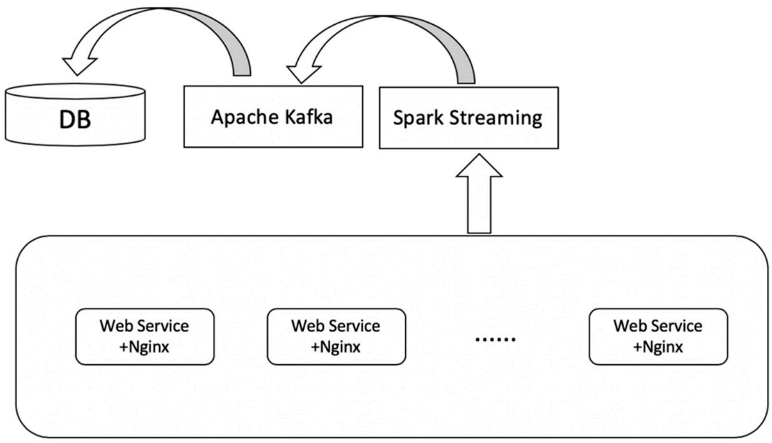

Figure 3. Real-time data process system.

In service-oriented architecture (SOA), to estimate the characteristics of dynamic and self-adaptability of web service composition, the authors used a Markov model to construct the evaluation model based on the state of software and state transition to describe the activity of software. Web service composition can be considered to be a task flow from an initial state to a final state. Corresponding to each task node, the specific web service is selected to be its state. If the currently selected web service corresponding to the task node is successfully invoked, the process continues to the next one, or it selects another task node, until it is successfully invoked. In this paper, the selection of a web service denotes the active state, and the currently selected web service depends on the current state and not on previous behavior. This successfully reflects the Markov behavior—the user invokes each subsequent web service based on the corresponding transition probability.

Figure 4 depicts the Markov model of the web service composition illustrated in Figure 2. Assuming the failure state to be F in which case each web service probably fails, the web service with failure will enter in state F. In addition, S denotes the final state and the specific web service w_{ij}

denotes a basic logical unit. Therefore, a series of web service flows can be established according to the composition.

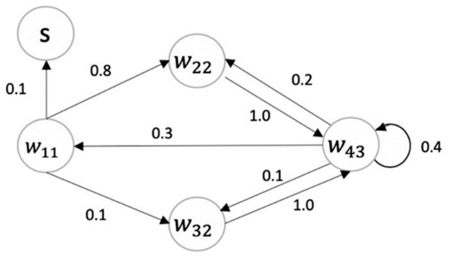

Figure 4. Web service composition Markov chain.

Based on the information recorded by the data processing system, the transition probability p_{ij} of the transition from the web service i to the web service j can be calculated, which is expressed by the values over the arcs in Figure 4. The collection of $p_{ij} \in [0, 1]$ describes the state transition relations between web services. The transition probability is given by

$$p_{ij} = \sum_{k=1}^{s} p_{sk} \times [\frac{\left|interact\left(c_i, c_j\right)\right|}{\left|interact\left(c_i, c_j\right)\right|_{i=1,2,\ldots,N}}]_{c_i, c_j \in S_k}$$

where S denotes the number of web service composition flows, p_{S_k} denotes the transition probability of the web service flow S_k in the web service composition, N denotes the number of web services, and $interact\left(c_i, c_j\right)$ denotes the number of interactions of web service i with web service j in the web service composition flow, S_k.

Therefore, in the case of a web service competition with n web services, the transition matrix of web service can be expressed as follows.

$$p = \begin{bmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{m1} & \cdots & p_{mn} \end{bmatrix} \tag{1}$$

Let us assume that the web service composition includes n web services and that web service i denotes the initial state. During successful operation of the web service composition, the web service i is transformed into its final state S with probability p_{iS}. The transition probability is given by the following equation:

$$p_{is} + \sum\nolimits_{j=1}^{n} p_{ij} = 1 \tag{2}$$

Equation (2) is derived from the web service composition in the absence of any failures. However, in practice, failure of web service compositions is common. Let r_i denote the reliability of the web service. Thus, when the web service i is invoked, its probability of failure is $1 - r_i$. In the Markov model of web service composition with $n + 2$ states, the transition matrix in the existence of failures is given by

$$\begin{cases} q_{ij} & = & r_i * p_{ij}, i = 1, 2, \ldots, n; j = 1, 2, \ldots, n \\ q_{iF} & = & 1 - r_i, i = 1, 2, \ldots, n \\ q_{SS} & = & q_{FF} = 1 \\ q_{ij} & = & 0, others \end{cases} \tag{3}$$

The transition matrix of a web service composition with n web services can be expressed as follows.

$$Q = \begin{bmatrix} \hat{P} & C \\ 0 & I \end{bmatrix}_{(n+2)\times(n+2)} \tag{4}$$

where $\hat{P}$ is an $n \times n$ matrix, $\left[q_{ij}\right]_{n\times n}$, denoting the single-step control transition probability between n web services, and C is an $n \times 2$ matrix denoting the control transition probability between the absorption states, S and F. In the web service composition depicted in Figure 4, the control transition matrix is given by

$$Q = \begin{array}{cc} \begin{bmatrix} 0 & p_{12}r_1 & p_{13}r_1 & 0 & p_{1s}r_1 & 1-r_1 \\ 0 & 0 & 0 & r_2 & 0 & 1-r_2 \\ 0 & 0 & 0 & r_3 & 0 & 1-r_3 \\ p_{41}r_4 & p_{42}r_4 & p_{43}r_4 & p_{44}r_4 & 0 & 1-r_4 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} & \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ S \\ F \end{array} \\ \begin{array}{cccccc} 1 & 2 & 3 & 4 & S & F \end{array} & \end{array} \tag{5}$$

This yields the following equation expressing the relationship between the reliability, R_s, of the web service composition and the reliability, $\mathrm{p_{is}}$, of each web service.

$$R_s = \sum\nolimits_{i=1}^{n} \left(I_n - \hat{Q}\right)_{1i}^{-1} r_i p_{is} \tag{6}$$

where I_n denotes the identity matrix of order n, and $\hat{Q}$ denotes the no-absorption transition probability matrix of order n. Using this equation, the reliability information of component levels can be fully utilized to calculate the reliability of the entire web service composition.

2.2. Reliability Assessment Method Based on Bayes Minimum Risk Decision

2.2.1. Reliability Assessment Scenario

In this paper, the authors performed reliability assessment based on Bayesian reliability assessment [22], by focusing on the problem of accurately and objectively verifying software reliability requirements. Bayesian reliability assessment is a statistical theory based on the Bayesian interpretation of probability. The probability of an event represents the degree of belief that it will occur, and prior information is used to predict the posterior distribution of reliability. Consider the parameter, θ, to be a random variable. The existing information was used to obtain prior distribution, the sample $X = (x_1, x_2, \ldots x_n)$, the prior information and current information was integrated to calculate the posterior distribution and statistically infer the value of θ. The mathematical model can be expressed by $f(x|\theta) \propto L(\theta|x)\pi(\theta)$, where $\pi(\theta)$ denotes the prior probability density function of θ, $f(x|\theta)$ denotes the posterior probability density function, and $L(\theta|x)$ denotes the likelihood function.

The authors proposed a two-stage reliability assessment of web service compositions—reliability growth stage and reliability demonstration stage. Assuming the reliability requirement to be (p_0, c), the probability of x failures in n test cases follows the binomial distribution—i.e., the probability of failure of the web service composition is p_s, which follows $Beta(a,b)$. Next, based on the aforementioned information obtained via reliability growth assessment, the distribution of the software failure probability, p, during reliability demonstration assessment was given by binary experiment—acceptable or rejected—the binomial sampling model with its conjugate beta distribution was widely used. The choice of the beta distribution as a prior is useful for several reasons: it is flexible enough to describe a variety of prior beliefs, it ensures that reliability is between (0, 1), and it is the conjugate prior for the binomial distribution [23,24]. The prior probability density of p_s can be expressed as follows:

$$\pi(p_s) = \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)} \tag{7}$$

where, $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx$.

During the reliability growth stage, the prior probability density p_s is confirmed and $\hat{p}_s$ (the estimate of p_s) is compared with the reliability requirement p_0—if $\hat{p}_s \le p_0$, reliability demonstration testing is conducted in the next step, if $\hat{p}_s > p_0$, need to repair, reliability growth testing is conducted again. Assuming that there are x failures in n test cases, the posterior probability density of p_s of the web service composition is given by:

$$\begin{aligned} f(p\big|x,n,a,b) &= Be(a+x,b+N-x) \\ &= \tfrac{p^{a+x-1}(1-p)^{b+n-x-1}}{B(a+x,b+n-x)} \end{aligned} \tag{8}$$

The expectation of the posterior probability density, p_s, is:

$$\begin{aligned} \hat{p}'_s &= E(p_s\big|x,n,a_0,b_0)\int_0^1 p_s B(a_0+x,b_0+n-x)dp_s \\ &= \tfrac{a_0+x}{b_0+n-x} \end{aligned} \tag{9}$$

The posterior expected value of p_s is known to be the minimum secondary loss estimate of the software failure reliability. It is used as the estimated value of the current probability of the web service composition. If $\hat{p}'_s < p_0$, the current reliability of the web service composition satisfies the reliability requirement, and the reliability growth test is deemed to be complete.

During the reliability demonstration stage, the posterior distribution p_s of the probability of failure during the growth test can be utilized, which integrates the information with the reliabilities of the web services and the web service composition. Thus, the distribution of failure probabilities can be obtained based on the prior information in the demonstration stage. The number of test cases that satisfy the reliability requirement, (p_0, c), is taken to be the minimum value of n in Equation (10):

$$\begin{aligned} P(p_s \le p_0) &= \int_0^{p_0} f(p\big|x,n,a,b)dp \\ &= \int_0^{p_0} \tfrac{p^{a+x-1}(1-p)^{b+n-x-1}}{B(a+x,b+n-x)}dp \\ &\ge \mathrm{c} \end{aligned} \tag{10}$$

Furthermore, by (10), after n test cases during which no failure is observed in reliability demonstration assessment, the reliability of web service composition satisfies the reliability requirement and can be accepted.

2.2.2. Bayes Decision

Reliability assessment was used to verify whether the software reliability satisfies the pre-defined reliability requirement, and the test result was assumed to be binary—acceptable or rejected. This transforms it into a decision-making problem (hypothesis testing) [23]. Let H_0 and H_1 denote the two decision states. H_0 denotes the test result meets reliability requirement, H_1 denotes the test result does not meet the reliability requirement. The result of the software reliability assessment can be described by

$$H_0 : \theta \le \theta_0,\ H_1 : \theta > \theta_0,$$

where θ denotes the actual failure rate, θ_0 denotes the failure rate of the reliability requirement, and the test significance level is taken to be c $(0 < c < 1)$. In this decision-making problem, the authors utilized the Bayesian decision function to describe the satisfaction of the reliability requirement by the assessed software in the presence of a certain number of failures. The decision function is given by

$$\delta(x) = P(H_0\big|X = x)$$

where x denotes the number of failures during reliability assessment based on Bayes' concept of assessment. The actual failure rate, θ, of the web service composition is taken to be a random variable, in this paper, assuming the prior distribution of parameter θ to be $G(\theta)$; therefore, the Bayesian decision function is given by

$$\delta_G(x) = \begin{cases} 1, & \theta \le \theta_0 \\ 0, & \theta > \theta_0 \end{cases}$$

according to (8), assuming the hyperparameter a and b are known, after computing. When the number of failures allowed by reliability requirement are 0, 1, and 2, N_0, N_1, and N_2 denote the upper limit number of test cases that need to be running, respectively. $\delta_G(x)$ can be transformed into the following (11):

$$\delta_G(x) = \begin{cases} 1, & x = 0 \; and \; n < N_0, x = 1 \; and \; n < N_1, x = 2 \; and \; n < N_2 \ldots \\ & 0, \; \theta > \theta_0 \end{cases} \tag{11}$$

The authors considered the result of reliability testing as Bayes decision problem; once the test result is wrong, which is bias from truth, it will cause the loss. In this paper, the authors utilized the Bayes decision idea in reliability testing of web service composition. In order to reduce the loss and improve accuracy of reliability testing, and control the bias of prior distribution $G(\theta)$ of parameter θ, in the next chapter, the authors propose the method of minimum risk (MMR) to produce the result with the lowest risk of decision.

2.2.3. Bayes Prior Hyperparameter Solving Method Based on MMR

In Bayesian reliability assessment, prior information is used to capture existing knowledge about unknown parameters. The efficiency and accuracy of the assessment is determined in terms of the prior probability distribution. An accurate prior distribution dramatically improves the required number of test cases [24]. The parameters of the prior distribution are referred to as hyperparameters, such as by a and b in (7). The values of the hyperparameters determine the accuracy of the prior probability distribution, which is an important factor in Bayesian decision-making [25]. A fundamental shortcoming of reliability assessment methods of web service compositions is that due to the accumulation of a large number of failures at each web service testing phase and a small number of failures in the web service composition testing phase, the lack of prior information makes the hyperparameters inaccurate and increases the risk of wrong decisions. In this section, the authors present a method for calculating a priori parameters based on an MMR, to minimize the risk of wrong Bayes decisions during reliability testing. The proposed method utilizes module-level prior information alongside the failure data of the constituent web services.

In Section 2.2, the authors consider the reliability assessment of web services as a decision-making problem; if wrong testing happened, incorrect assessments induce the loss of accuracy. The authors used the loss function to describe the risk (12), which represents the loss induced by wrong assessment. To measure the loss accurately, the authors used the risk function to express the expected value of the loss. Consequently, the decision-making problem can be converted into that of identifying a method that minimizes the risk during testing. The loss function can be indicated in following:

$$\begin{aligned} L_0 = (\theta, d_0) = a(\theta - \theta_0) I(\theta > \theta_0) \\ L_1(\theta, d_1) = a(\theta - \theta_0) I(\theta \le \theta_0) \end{aligned} \tag{12}$$

where a denotes a positive constant and $d = \{d_0, d_1\}$ indicates the test results, so according to result of reliability testing, the decision rule is: d_0 denotes that accepted H_0, d_1 denotes that rejected H_0. $I(\theta > \theta_0)$ represents the actual failure probability of the web service composition being larger than the reliability requirement. On the other hand, $I(\theta < \theta_0)$ represents the actual failure probability of web service composition being smaller than the reliability requirement. $L_0 = (\theta, d_0)$ denotes the custom risk corresponding to cases in which the software is accepted despite the failure probability being larger than the reliability requirement and $L_1 = (\theta, d_1)$ denotes the producer risk corresponding to cases in which the software is rejected despite the failure probability being smaller than the reliability requirement.

If the sample set is denoted by $\Omega = \{x|x\rangle 0\}$, the parameter set is denoted by $\Theta = \left\{\theta > 0\middle|\int_\Omega f(x|\theta)dx = 1\right\}$, then, corresponding to the prior distribution $G(\theta)$, the risk function of the decision function, $\delta(x)$, is given by

$$\begin{aligned} R(\delta(x), G(\theta)) &= E_{(x,\theta)}[L(\theta, \delta(x))] \\ &= \int_\Theta \left[\int_\Omega L(\theta, \delta) f(x|\theta)dx\right] dG(\theta) \end{aligned} \tag{13}$$

substituting (12) into (13) yields

$$\begin{aligned} R(\delta(x), G(\theta)) &= \int_\Theta \int_\Omega [L_0(\theta, d_0) f(x|\theta)\delta(x) + L_1(\theta, d_1) f(x|\theta)(1 - \delta(x))] dx dG(\theta) \\ &= \int_\Theta \int_\Omega [a(\theta - \theta_0)(1 - I(\theta < \theta_0))\delta(x) + L_1(\theta, d_1)(1 - \delta(x))] f(x|\theta) dx dG(\theta) \\ &= \int_\Theta \int_\Omega [a(\theta - \theta_0)\delta(x) + L_1(\theta, d_1)] f(x|\theta) dx dG(\theta) \end{aligned} \tag{14}$$

and further substituting (11) and $\int_\Omega f(x\big|\theta)dx = 1$ into (14) yields

$$R(\delta(x), G(\theta)) = a \int_{x \in \{x|\theta \le \theta_0\}} \beta(x) dx + C_G \tag{15}$$

where

$$\begin{aligned} \beta(x) &= \int_\Omega^{\theta_0} (\theta - \theta_0) f(x|\theta)\pi(\theta) d\theta \\ C_G &= \int_\Theta L_1(\theta, d_1)\pi(\theta) d\theta \\ &= a \int_0^{\theta_0} (\theta_0 - \theta)\pi(\theta) d\theta \end{aligned}$$

Therefore, during reliability assessment of web services, substituting the distribution function, $f(x\big|\theta)$, of the number of failures and the prior probability density function, $\pi(\theta)$, of failure probabilities into (15), the risk function of the decision function, $\delta(x)$, can be obtained as the following.

$$\begin{aligned} R = a \int_0^e \int_0^{p_0} (p - p_0) C_N^x p^x (1-p)^{N-x} \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)} dp dx \\ + a \int_0^{p_0} (p - p_0) \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)} dp \end{aligned} \tag{16}$$

where e denotes the maximum number of permissible failures during the assessment of the web service composition. After obtaining the estimated values, $\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_n$, of the failure probabilities corresponding to all web services, the authors can utilize (6) to calculate the estimated value, p_s, of the failure probability of the web service composition as follows.

$$\hat{p}_s = F(\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_n) \tag{17}$$

The prior probability distribution of the failure probability, p_s, follows $Beta(a, b)$. Therefore, the expected value of the failure probability, p_s, of the web service composition is given by

$$\overline{p_s} = E(\pi) = \int_0^1 p_s Be(a, b) dp_s = \frac{a}{a + b} \tag{18}$$

The expected value of $\mathrm{p_s}$ is taken to be the estimated value of the failure probability of the web service composition, yielding

$$\frac{a}{a + b} = F(\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_n) \tag{19}$$

Therefore, the problem of calculating prior hyperparameters, a and b, can be transformed into the problem of identifying the optimal solutions of (20) while satisfying the MMR.

$$Min\{R\} = Min\Big\{a\int_0^e\int_0^{P_0}(p-p_0)C_N^x p^x(1-p)^{N-x}\frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}dpdx + a\int_0^{P_0}(p-p_0)\frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}dp\Big\} \quad \frac{a}{a+b} = F(\hat{p}_1,\hat{p}_2,\ldots,\hat{p}_n)a>0,b>0 \tag{20}$$

where x denotes the permissible number of failures and N denotes the total number of test cases. Finally, based on the values of the hyperparameters, a and b, the failure probability distribution can be obtained, and the number of test cases, n, in (10) satisfies the reliability requirement of the web service composition assessment.

3. Results and Discussion

The authors constructed a controlled environment comprising a set of web services and deployed four web services in the Amazon web service (AWS). Each service was executed on an Amazon machine image equipped with CentOS 8 (x86_64), ASP.NET Core [26]. The architecture of the web service composition in this paper is illustrated in Figure 5. It includes four web services, and the arrows in the figure denote their invocations. The authors utilized the tool, SoapUI, to generate the soap requests according to web service description language (WSDL) [27] to simulate the process of user-invoked web services.

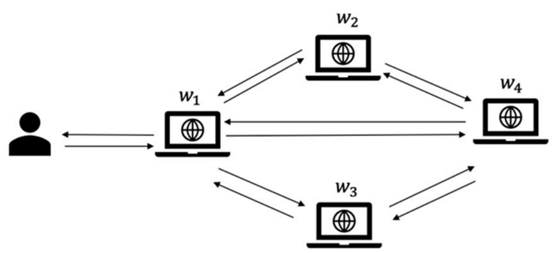

Figure 5. Web service composition with invocations.

3.1. Reliability Assessment of Web Service Composition

In this paper, the authors primarily considered the errors induced by web service responses. To this end, the authors utilized a very useful Linux utility called traffic control (tc) to assign a threshold of 10 s for the response time to adequately simulate and control the network latency. If a response is so slow that its delay is noticeable by a customer (i.e., 20% or more), then the service is deemed to have effectively failed.

The reliability requirement of the web service composition set by the developer is $p = 0.004$. Reliability assessment of the web services revealed the corresponding failure probabilities, as presented in Table 2—$\hat{p}_1 = 0.081$, $\hat{p}_2 = 0.024$, $\hat{p}_3 = 0.096$ and $\hat{p}_4 = 0.038$. By (6), the failure probability, p_s, of the web service composition can, therefore, be calculated as follows.

$$\hat{p}_s = F(\hat{p}_1,\hat{p}_2,\ldots,\hat{p}_4) = 0.0081 \tag{21}$$

Table 2. Failure probabilities of web services.

Web Service	Failure Probability
w_1	0.081
w_2	0.024
w_3	0.096
w_4	0.038

Substituting (21) into (19) yields

$$\frac{a}{a+b} = 0.0081 \tag{22}$$

Using (20), the values of the hyperparameters were calculated to be a = 1 and b = 122.5 following MMR. Then, reliability assessment was initiated with reliability growth assessment. By (21), the reliability, $\hat{p}_s$, of the web service composition was calculated to be 0.0081 and, as 0.0081 > 0.004, the reliability of web service composition was deemed to not satisfy the pre-defined reliability requirement, after repairing. Thus, reliability growth assessment was continued. When performing the 250th test case, the failure occurred, then correcting it. By (8), the failure probability can be calculated as follows.

$$\hat{p}_s' = \frac{a_0 + x}{a_0 + b_0 + n} = \frac{1+1}{1 + 122.5 + 250} = 0.0054 > 0.004$$

Evidently, the current reliability of web service composition still did not satisfy the reliability requirement. Thus, testing was continued until the 1529th test case, when the 6th failure was observed. The corresponding failure probability was calculated as follows.

$$\hat{p}_s' = \frac{a_0 + x}{a_0 + b_0 + n} = \frac{1+6}{1 + 122.5 + 1829} = 0.0036 < 0.004$$

At this point, the reliability of the web service composition finally satisfied the reliability requirement, and therefore, the software reliability growth assessment was deemed to be complete. Meanwhile, the values of the hyperparameters of the posterior distribution of the failure probability p were calculated to be $a_0 = 7,$ and $b_0 = 1760.1$.

$$f(p_s) = Beta(7,\ 1760.1) = \frac{p_s^6 (1-p_s)^{1759.1}}{B(7,\ 1760.1)}$$

To verify whether the reliability of the web service composition satisfied the reliability requirement, the following equation was used, following (20), to calculate the maximum number of test cases required for the reliability demonstration test.

$$\int_0^{p_0} \frac{p^{a_0+j-1}(1-p)^{b_0+N_{j+1}-j-1}}{B\left(a_0+j, b+N_{j+1}-j\right)} dp \geq c$$

Finally, the authors compared the required numbers of test cases using MMR, Bayesian Monte Carlo method (BMCM) [28], and expert scoring method (ESM) [29] to obtain a reliability demonstration requirement of $(p_0, c) = (0.005, 0.9)$. With an increasing number of failures, the maximum number of test cases corresponding to different methods was required to satisfy the requirements presented in Table 3.

Table 3. Number of reliability demonstration test cases with different methods.

Number of Failures	Number of Test Cases		
	BMCM	**ESM**	**MMR**
0	4602	3932	3692
1	7635	5478	4573
2	9402	7123	6432
3	10,041	8232	8923
4	11,900	10,345	10,401
5	16,104	13,335	11,452

In the case of BMCM, the unavailability of prior information during the first execution of the web service composition induced a uniform distribution of the hyperparameters over the domain. In this case, $Beta$(2, 2) was adopted to represent the lack of prior knowledge, which assigned the prior reliability centers around 0.5, with equal probabilities of failure and success. Then, the Bayes formula based on the prior distribution was used to calculate the n test cases. In a beta-binomial conjugation framework, it is useful to think that the hyperparameters of the prior distribution (beta distribution) of system failure probability correspond to a certain number of pseudo observations with the properties specified by these hyperparameters [30,31].

In the case of ESM, let us assume that there are n experts who assign judgment intervals of $[\theta_L, \theta_H]$ corresponding to the failure probability, θ_k, associated to the kth stage of growth assessment, and that $\overline{\theta}_L$ and $\overline{\theta}_H$ denote the means of θ_L and θ_H, respectively, over all experts. Then, the expectation, E, and the variance, V, can be calculated using the following equations.

$$E = \frac{\overline{\theta}_L + \overline{\theta}_H}{2}$$
$$V = \frac{\left(\overline{\theta}_H - \overline{\theta}_L\right)^2}{12}$$

The hyperparameters, a and b, can be evaluated using the following equations.

$$a = \frac{E \cdot (V - nE)}{D}$$
$$b = \frac{(n-E)(V-nE)}{D}$$

where $D = (n-1)E^2 + n(E - V)$.

The judgment interval, $[\theta_L, \theta_H]$, of each expert reflects their judgement as elicited via interviews. Following the completion of the three steps of reliability growth assessment presented in Table 4, the posterior distribution was obtained to be Beta (61.27,324.92).

Table 4. Information of experts.

Stage in Growth Testing	Expert 1	Expert 2	Expert 3	Expert 4	Beta Distribution		Hyperparameter	
					Expected	Variance	a	b
1	(0.003, 0.0072)	(0.0035, 0.0075)	(0.0040, 0.0074)	(0.0038, 0.0072)	0.0054	1.17×10^{-6}	17.58	140.41
2	(0.0080, 0.0088)	(0.0060, 0.0086)	(0.0065, 0.0090)	(0.0063, 0.0085)	0.0077	3.42×10^{-6}	21.22	270.78
3	(0.0075, 0.0095)	(0.0077, 0.0093)	(0.0076, 0.0092)	(0.0078, 0.95)	0.0085	2.48×10^{-6}	33.27	311.92

The numbers of required test cases corresponding to different methods are presented in Table 3. The required number of test cases in the case of BMCM is $n = 4602$, which implies that if no failure is observed in 4602 successive test cases, the reliability of web service composition is deemed to satisfy the reliability requirement, $(p_0, c) = (0.005, 0.9)$, and the software can be accepted. However, if a failure is noticed in, say, the 3021st test case, the assessment must be continued for 4614 (7635 − 3021 = 4614) further test cases. If no failure is observed during these iterations, the web service composition is deemed to satisfy the reliability demonstration requirement. However, if an additional failure was noticed during this stage as well at, say, the 5500th test case, the assessment is to be continued for another additional 3902 (9402 − 5500 = 3902) test cases. The satisfaction of the reliability demonstration requirement is then adjudged once again. The process continues in an analogous fashion.

The experimental data indicated that the number of test cases required by MMR was significantly lower than that required by BMCM and ESM. Corresponding to 0 failures, the numbers of test cases required by BMCM and ESM were 4602 and 3932, respectively, whereas that required by MMR was merely 3692. This represents reductions of 910 and 670, respectively, corresponding to reduction rates of 19.8% and 14.6%, respectively. Furthermore, in the case of MMR, the required number of test cases exhibited the smallest increment when the number of failures was increased from 0 to 1—for BMCM, ESM, and MMR, the required numbers of test cases increased by 65.9%, 39.3%, and 23.9%, respectively, in this case. Similarly, when the number of failures was increased from 4 to 5, the increase in the numbers of test cases required by BMCM, ESM, and MMR were observed to be 35.3%, 28.9%, and 10.1%, respectively.

Thus, the experiments demonstrated that in the case of identical prior information, MMR was the most accurate method to calculate the prior distribution, and it significantly reduced the number of test cases required during reliability demonstration assessment. In the case of ESM, the opinions of experts made the value of hyperparameters more subjective, which wielded a greater influence on the postdistribution. In the case of BMCM, the lack of prior information induced a uniform initial distribution of the hyperparameter over the domain. Subsequently, the distributions were repeatedly sampled to obtain the expected values of the hyperparameters. This procedure usually involves complex numerical calculations and is difficult to implement.

3.2. Goodness-of-Fit of the Distribution

In this subsection, the authors analyze the goodness-of-fit of the distribution to the truth data; maximum likelihood estimation (MLE) was used to evaluate the accuracies of the corresponding distributions obtained using hyperparameters estimated using BMCM, ESM, and MMR. MLE is a method to evaluate models using specific data—given different prior to distributions; it determines the corresponding likelihoods of the actual sample to determine the confidences of the prior distributions. The likelihood is termed the confidence factor; and it is proportional to the proximity of the prior distribution to the ground truth. Obviously, the distribution that is closest to the ground truth is deemed to be the best.

MLE was conducted using the following procedure. Let us assume that the failure probability distribution of the web service composition, $X = \{p_1, p_2 \ldots p_n\}$, is given by $f(X|\theta)$, where θ denotes the hyperparameters, a and b, and the density function of the prior distribution of θ is $\pi(\theta)$. After obtaining X, the density function $\pi(\theta|X)$ of the posterior distribution can be expressed as

$$\pi(\theta|X) = \frac{f(X|\theta)\pi(\theta)}{\int_\theta f(X|\theta)\pi(\theta)d_\theta} \tag{23}$$

where $f(X|\theta)$ reflects the knowledge of θ based on the observed X. If there are two parameters, θ_1 and θ_2, the authors obtain

$$f(X|\theta_1) > f(X|\theta_2) \tag{24}$$

From the perspective of statistics, the probability of observed X with the parameter θ_1 is higher than X with θ_2. Thus, (24) can be interpreted as the likelihood of $\theta = \theta_1$ being larger than that of $\theta = \theta_2$. Thus, when X is constant, $f(X|\theta)$ is independent of the prior distribution of θ as the function with θ; instead, it merely reflects the characteristics of the sample.

The denominator of (23) is given by

$$m(X) = \int_\theta f(X|\theta)\pi(\theta)d_\theta \tag{25}$$

This is the marginal distribution of X that depends on $f(X|\theta)$. $f(X|\theta)$ is the conditional distribution density and, for a given θ, the marginal distribution describes the action of the sample X under the prior distribution, $\pi(\theta)$, and the sample follows this distribution. If $\pi(\theta)$ follows a certain distribution, the parameter of $\pi(\theta)$ is obtained from the maximum value of $m(X)$—i.e., the MLE.

If there are m prior distributions that follow $f(X|\theta)$, the failure probability samples are $p_1, p_2 \ldots p_n$. Let $\pi_i(\theta)$ be the i^{th} prior distribution based on the sample. If the ground truth prior distribution is denoted by $\pi(\theta)$, the m prior distributions can be integrated into the ground truth prior distribution using the following equation.

$$\pi(\theta) = \sum_{i=1}^{m} \varepsilon_i \pi_i(\theta) \tag{26}$$

where ε_i denotes the weight of ith prior distribution, $\pi_i(\theta)$, and $\sum_{i=1}^{m} \varepsilon_i = 1$.

The marginal distribution of the prior distribution, $\pi_i(\theta)$, is given by

$$m(x|\pi_i) = \int_\theta f(X|\theta)\pi_i(\theta)d_\theta \quad (27)$$

The marginal distribution of the ground truth prior distribution, $\pi(\theta)$, is given by

$$m(x|\pi) = \int_\theta f\left(X\middle|\theta\right)\pi(\theta)d_\theta \quad (28)$$

The failure data, $p_1, p_2 \ldots p_n$, generated by the marginal distribution, $m(x|\pi_i)$, can be used to obtain the following likelihood function.

$$L(X|\pi_i) = \prod_{i=1}^{n} m(x_i|\pi_i) \quad (29)$$

By MLE, $L(X|\pi_i)$ is bigger to represent the closer proximity to the ground truth, the weight of $\pi_i(\theta)$ is bigger, high probability to choose this prior distribution. The confidence factor can be obtained as follows.

$$\varepsilon_k = \frac{L(X|\pi_i)}{\sum_{i=1}^{i} L(X|\pi_i)} \quad (30)$$

The confidence factor represents the probability of a prior distribution being close to the truth distribution. Thus, if the confidence factor of a distribution is high, it is highly feasible to choose it as a prior distribution.

Next, the authors compared the accuracies of the distribution constructed using the hyperparameters estimated using BMCM, ESM, and MMR. The failure probability follows $Beta(a, b)$, where a and b are hyperparameters.

First, certain defects were introduced into the web services, following the discussion presented in [32]. The web service composition considered in this section is depicted in Figure 4. The actual failure probability following the introduction of the defects was regarded as the ground truth, the failure probabilities obtained via various method were regarded as evaluation values, X denotes the ground truth data, and MLE was used to determine the distribution that is closest to the truth. The failure probabilities of each web service obtained via module-level testing are presented in Table 5—$\hat{p}_1 = 0.011$, $\hat{p}_2 = 0.034$, $\hat{p}_3 = 0.096$ and $\hat{p}_4 = 0.018$. Based on injected defects, the ground truth of the failure probability of the web service composition was calculated to be $X = (0.184,\ 0.299,\ 0.156)$, which follows $Beta(3.0,\ 6.0)$.

Table 5. Failure probabilities of web services.

Web Services	Failure Probability
w_1	0.011
w_2	0.034
w_3	0.096
w_4	0.018

Then, the authors calculated the failure probabilities using the predictions obtained from BMCM, ESM, and MMR as evaluation values. In the case of BMCM, the initial non-informative prior distribution was assumed to be $Beta(2,\ 2)$. In the case of ESM, after assessment over 10 h, the prior distribution of the web service composition was estimated to by expert interview and the initial prior distribution was $Beta(3,\ 6.3)$. In the case of MMR, by (20) and (22), the initial prior distribution was $Beta(2.5,\ 3.9)$.

After three steps of growth assessment, the probability distributions corresponding to the three methods were observed to be $\pi_1 = Beta(3.1,\ 5.9)$, $\pi_2 = Beta(2.7,\ 6.6)$, and $\pi_3 = Beta(3.5,\ 7)$ (π_1 denotes MMR, π_2 denotes BMCM, and π_3 denotes ESM).

By (27), the marginal distribution functions are given by

$$\begin{aligned} m(x|\pi_1) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1} = \frac{\Gamma(9)}{\Gamma(3.1)\Gamma(5.9)}x^{2.1}(1-x)^{4.9} \\ m(x|\pi_2) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1} = \frac{\Gamma(9.3)}{\Gamma(2.7)\Gamma(6.6)}x^{1.7}(1-x)^{5.6} \\ m(x|\pi_3) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1} = \frac{\Gamma(10.5)}{\Gamma(3.5)\Gamma(7)}x^{2.5}(1-x)^{6} \end{aligned}$$

Then, by (29), the likelihood functions are given by

$$\begin{aligned} L(X|\pi_1) &= 8.3003e-005 \\ L(X|\pi_2) &= 4.0761e-005 \\ L(X|\pi_3) &= 1.2588e-005 \end{aligned}$$

By (30), the respective confidences are given by

$$\begin{aligned} \varepsilon_1 &= 0.6707 \\ \varepsilon_2 &= 0.3294 \\ \varepsilon_3 &= 0.1321 \end{aligned}$$

As to $\varepsilon_1 > \varepsilon_2 > \varepsilon_3$, the prior distribution corresponding to MMR is closest to the ground truth and is, therefore, the best candidate for the prior distribution.

The curves of *Beta* corresponding to the three different methods are illustrated in Figure 6. Line 0 corresponds to the ground truth, Line 1 corresponds to the prior distribution, $\pi_1 = Beta(3.1,\ 5.9)$, Line 2 corresponds to the prior distribution, $\pi_2 = Beta(2.7,\ 6.6)$, and Line 3 corresponds to the prior distribution, $\pi_3 = Beta(3.5,\ 7)$. Line 1, corresponding to MMR, is evidently a significantly better fit compared to the other curves. Therefore, calculating the hyperparameters using MMR is the most reasonable approach in this case.

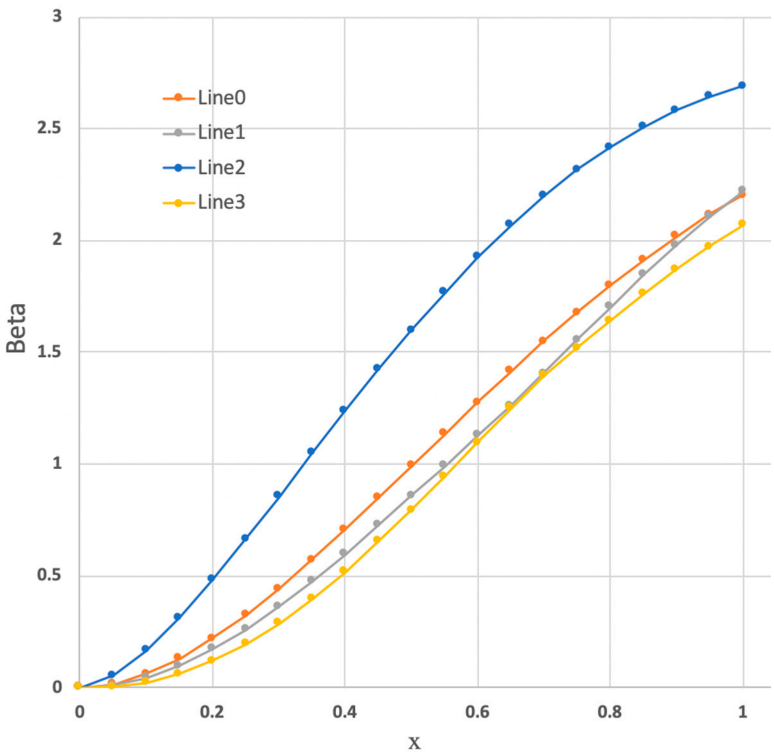

Figure 6. Comparison of the various fits.

The accuracy of reliability estimation depends on the prior distribution. Based on the aforementioned discussion, the prior distribution constructed using hyperparameters calculated using MMR is the closest to the ground truth distribution. Thus, MMR can be used efficiently utilize the prior information of web services, and its results exhibit the lowest decision risk and are more objective and consistent with the actual situation.

4. Conclusions

In this paper, a reliability assessment method was proposed for web service compositions. To this end, the authors used transfer probabilities to represent the inter-relationships between the constituent web services, and used a Markov model to transfer the probability matrix and calculate the reliability of web service composition. The proposed method makes full use of failures data of the web services, thereby adequately resolving the problem of a lack of prior information in web service compositions.

The authors divided the reliability assessment procedure into two stages—the reliability growth stage and the reliability demonstration stage. The primary problems faced during the reliability growth testing are inaccurate prior distributions and the requirement of large numbers of test cases, which increases the test duration. The authors used Bayes reliability assessment for reliability testing, thereby fully utilizing prior information of the web services, and proposed reliability demonstration assessment based on prior information. The authors assumed the number of failures, x, to follow a binomial distribution, $Beta(a, b)$. The authors compared the estimated failure probabilities, p_s, to p_0 and formulated a completion condition for the reliability growth assessment. During the reliability demonstration stage, the authors considered the reliability assessment result as essentially a Bayesian decision-making (hypothesis test) problem. During reliability assessment, prior hyperparameters affect the accuracy of the prior probability distribution and play an important role in the Bayesian decision. The authors discussed the importance of hyperparameters in the Bayesian decision, and proposed the MMR method to calculate prior hyperparameters. Experiments were conducted to demonstrate that the proposed method exhibited a smaller number of required test cases compared to two other methods. This increases the efficiency of reliability demonstration assessment and avoids the adoption of rash inferences about the reliability of web service compositions.

In future research, the authors intend to consider web service architectures of more complicated designs that require significantly more calculations, considering different reliability growth models; the efficiency of the proposed method should also be improved. In addition, reliability assessment based on MMR should be repeated to further improve the proposed method. The final aim is to efficiently utilize available resources during reliability assessment of web service compositions.

Author Contributions: Writing—original draft, Y.S.; Writing—review & editing, Y.W. and D.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China grant number U1736110.

Acknowledgments: My deepest gratitude goes first and foremost to Yawen Wang and Dahai Jin, my supervisor, for constant encouragement and guidance. He walked me through all the stages of the writing of this thesis. Without his consistent and inspiring instruction, this thesis could not have reached its present form. My thanks also go to my beloved family for their loving consideration and great confidence in me throughout the years, and to my friends and fellow classmates, who gave up their time to help me work out any issues during the difficult course of this thesis.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

Some abbreviations and notations are listed below:

MMR	Method of minimum risk;
BMCM	Bayesian Monte Carlo method;
ESM	Expert scoring method;
SOA	Service-oriented architecture;
QoS	Quality of service;
AWS	Amazon web service;
WSDL	Web service description language;
Tc	Traffic control;
BMCM	Bayesian Monte Carlo method;
ESM	Expert scoring method;
W_i	Abstract web service candidates;
w	A specific web service;
t	Duration;
C_n	The numbers of invocations of the web service;
n	Failures of the web service;
$F(w)$	The probability of failure of the web service w;
$r(w)$	Reliability of the web service w;
p_{ij}	The state transition relations between web services;
N	The number of web services;
$interact\left(c_i, c_j\right)$	The number of interactions of web service i with web service j;
S_k	Service composition flow;
$\hat{P}$	$n \times n$ matrix;
I_n	The identity matrix of order n;
$\hat{Q}$	The no-absorption transition probability matrix of order n;
R_s	Reliability of web service composition;
θ	Parameter of distribution;
$\pi(\theta)$	The prior probability density function;
$f(x\|\theta)$	The posterior probability density function;
$L(\theta\|x)$	The likelihood function
(p_0, c)	Reliability requirement;
p_s	The probability of failure of the web service composition;
H_0	The test result meets reliability requirement;
H_1	Test result does not meet the reliability requirement;
$G(\theta)$	Prior distribution of parameter;
d_0	Accepted H_0;
d_1	Rejected H_0;
$I(\theta < \theta_0)$	The actual failure probability of web service composition being smaller than the reliability requirement;
$I(\theta > \theta_0)$	The actual failure probability of the web service composition being larger than the reliability requirement;
L_0	The custom risk corresponding to cases in which the software is accepted despite the failure probability being larger than the reliability requirement;
L_1	The producer risk corresponding to cases in which the software is rejected despite the failure probability being smaller than the reliability requirement;
$\delta(x)$	The risk function of the decision function;
$\pi(\theta)$	Prior probability density function;
N	The total number of test cases;
$a,\ b$	Hyperparameters;
E	Expectation of ESM;
V	Variance of ESM;
$m(x\|\pi_i)$	Marginal distribution of the prior distribution $\pi_i(\theta)$;
ε_k	Confidence factor.

References

1. Lee, C.H.; Kim, Y.-H.; Park, D.H. Evaluation of multi-tasking software system performance with consideration of module dependency. *J. Softw. Maint. Evol. Res. Pr.* **2011**, *23*, 361–374. [CrossRef]
2. Straub, D.; Papaioannou, I. Bayesian updating with structural reliability methods. *J. Eng. Mech.* **2015**, *141*, 04014134. [CrossRef]
3. Rausand, M.; Barros, A.; Hoyland, A. *System Reliability Theory: Models, Statistical Methods, and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2020.
4. Sinharay, S.; Stern, H.S. On the sensitivity of Bayes factors to the prior distributions. *Am. Stat.* **2002**, *56*, 196–201. [CrossRef]
5. Li, M.; Meeker, W.Q. Application of Bayesian methods in reliability data analyses. *J. Qual. Technol.* **2014**, *46*, 1–23.MLA. [CrossRef]
6. Anderson-Cook, C.M. Opportunities and issues in multiple data type meta-analyses. *Qual. Eng.* **2009**, *21*, 241253. [CrossRef]
7. Csenki, A. Bayes predictive analysis of a fundamental software reliability model. *IEEE Trans. Reliab.* **2002**, *39*, 177–183. [CrossRef]
8. Moala, F.A.; Rodrigues, J.; Tomazella, V.L.D. A note on the prior distributions of weibull parameters for the reliability function. *Commun. Stat. Theory Methods* **2009**, *38*, 1041–1054. [CrossRef]
9. Burgman, M.; McBride, M.; Ashton, R.; Speirs-Bridge, A.; Flander, L.; Wintle, B.; Fidler, F.; Rumpff, L.; Twardy, C. Expert status and performance. *PLoS ONE* **2011**, *6*, e22998. [CrossRef] [PubMed]
10. Hgan, A.O.; Buck, C.; Daneshkhah, A.; Eiser, J.; Garthwaite, P.; Jenkinson, D.; Oakley, J.; Rakow, T. *Uncertain Judgements: Eliciting Experts Probabilities*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
11. McConway, K. The Combination of Experts Opinions in Probability Assessment: Some Theoretical Considerations. Ph.D. Thesis, University College London, London, UK, 1978.
12. Li, Z.S.; Guo, J.; Xiao, N.-C.; Huang, W. Multiple priors integration for reliability estimation using the Bayesian melding method. In Proceedings of the 2017 Annual Reliability and Maintainability Symposium (RAMS), Orlando, FL, USA, 23–26 January 2017; pp. 1–6.
13. Pham, C.; Tran, N.H.; Ren, S.; Saad, W.; Hong, C.S. Traffic-Aware and Energy-Efficient vNF Placement for Service Chaining: Joint Sampling and Matching Approach. *IEEE Trans. Serv. Comput.* **2020**, *13*, 172–185. [CrossRef]
14. Di Mauro, M.; Longo, M.; Postiglione, F.; Tambasco, M. Availability modeling and evaluation of a network service deployed via NFV. In *Communications in Computer and Information Science*; Springer: Berlin/Heidelberg, Germany, 2017; Volume 766, pp. 31–44.
15. Bian, S.; Huang, X.; Shao, Z.; Gao, X.; Yang, Y. Service Chain Composition with Failures in NFV Systems: A Game-Theoretic Perspective. In Proceedings of the ICC 2019—2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019; pp. 1–6. [CrossRef]
16. di Mauro, M.; Longo, M.; Postiglione, F. Availability Evaluation of Multi-tenant Service Function Chaining Infrastructures by Multidimensional Universal Generating Function. *IEEE Trans. Serv. Comput.* **2018**, *1*. [CrossRef]
17. Fan, T.H.; Balakrishnan, N.; Chang, C. The Bayesian approach for highly reliable electro-explosive devices using one-shot device testing. *J. Stat. Comput. Simul.* **2009**, *79*, 1143–1154. [CrossRef]
18. Kuelbs, J.; Philipp, W. Almost sure invariance principles for partial sums of mixing B-valued random variables. *Ann. Probab.* **1980**, *8*, 1003–1036. [CrossRef]
19. Dai, Y.-S.; Xie, M.; Long, Q.; Ng, S.-H. Uncertainty analysis in software reliability modeling by bayesian analysis with maximum-entropy principle. *IEEE Trans. Softw. Eng.* **2007**, *33*, 781–795. [CrossRef]
20. Guan, Y.; Dy, J. Sparse probabilistic principal component analysis. *Artif. Intell. Stat.* **2009**, *5*, 185–192.
21. Bergman, S.; Schiffer, M. *Kernel Functions and Elliptic Differential Equations in Mathematical Physics*; Courier Corporation: Chelmsford, MA, USA, 2005.
22. Singpurwalla, N.D. *Reliability and Risk: A Bayesian Perspective*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
23. Sebald, D.J.; Bucklew, J.A. Support vector machines and the multiple hypothesis test problem. *IEEE Trans. Signal Process.* **2001**, *49*, 2865–2872. [CrossRef]

24. Rothermel, G.; Untch, R.H.; Chu, C.; Harrold, M.J. Prioritizing Test Cases for Regression Testing. 2001. Available online: https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1017&context=csearticles (accessed on 3 December 2020).
25. Lindley, D.V. The use of prior probability distributions in statistical inference and decision. In Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20–30 July 1960; pp. 453–468.
26. Sajee, M.; Varia, J. Overview of Amazon Web Services. 2014. Available online: https://docs.aws.amazon.com/whitepapers/latest/aws-overview/aws-overview.pdf (accessed on 3 December 2020).
27. Wikipedia. LoadUI. Available online: https://en.wikipedia.org/wiki/LoadUI (accessed on 3 December 2020).
28. Dai, Z.; Wang, Z.; Jiao, Y. Bayes Monte-Carlo assessment method of protection systems reliability based on small failure sample data. *IEEE Trans. Power Deliv.* **2014**, *29*, 1841–1848. [CrossRef]
29. Walls, L.; Quigley, J. Building prior distributions to support Bayesian reliability growth modelling using expert judgement. *Reliab. Eng. Syst. Saf.* **2001**, *74*, 117–128. [CrossRef]
30. Hamada, M.; Martz, H.F.; Reese, C.S.; Graves, T.; Johnson, V.; Wilson, A.G. A fully bayesian approach for combining multilevel failure information in fault tree quantification and optimal follow-on resource allocation. *Reliab. Eng. Syst. Saf.* **2004**, *86*, 297–305. [CrossRef]
31. Christensen, R.; Johnson, W.; Branscum, A.; Hanson, T.E. *Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians*; CRC Press: Boca Raton, FL, USA, 2011.
32. Wang, H.; Ouni, A.; Kessentini, M.; Maxim, B.; Grosky, W.I. Identification of web service refactoring opportunities as a multi-objective problem. In Proceedings of the 2016 IEEE International Conference on Web Services (ICWS), San Francisco, CA, USA, 27 June–2 July 2016; pp. 586–593.

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).