An Efficient Bayes Classifiers Algorithm for Traceability of Food Supply Chain Management using Internet of Things

S Balamurugan, A Ayyasamy, K Joseph

Abstract— The conventional Food Supply Chain Management (FSCM) faces a variety of provocation such as ambiguity, security, cost, complication and quality concerns. To resolve these issues, supply chain must be precise. A challenging assignment in today's food industry is distributing the high quality of foods throughout the supply chain management. In this paper, proposes an efficient Bayes Classifiers Algorithm which integrated with FSCM using Internet of Things (IoT) to allow tracking, tracing and managing the entire process of food supply chain such as supplier, exporter and customers. The objective of this paper is to determine the food safety and to optimize chronological data produced to analyze the effective possibility of future assump- tions. It also aims to foods carrying from the manufacturers to the customers with help of IoT technologies to bond the producer to the customer with delivery of high class of food products. IoT based Food Supply Chain Traceability is utilized t o data transaction effectively with indeterminate, uncertain and insufficient information. So the proposed efficient Bayes Classifiers Algorithm will be capable to overcome all provocation of conventional supply chain and afford secure background and food safety for FSCM process using IoT technology.

Index Terms—Bayes classifier algorithm, Tree Augmented Naive Bayes, Traceability, FSCM, Food safety and IoT

I. INTRODUCTION

Supply chain is a set of procedures and elements (provider, distributer, dealer, shopkeeper and consumer) which are interrelated to fulfil customer needs [1] . During the last decades, customer belief in the food manufacturing was very much damaged because of preserving the nutritional attribute [2] of products during transportation from producer to customer. Supply chain management signifies having the correct product in the correct quantity at the correct time at the correct set in the correct state to the correct consumer [3] . For successful system of supply chain, an intelligent technique plays a significant role [4] . It has capability to incorporate dif- ferent procedures, providers and consumers within and outside through improving communication, gathering and transport of data and information and then enhance the performance of supply chain [5] .One of the most significant improvements of intelligent technique is the Internet of Things (IoT).

  S Balamurugan*, Ph.D., Research Scholar, Department of CSE, Annamalai University, Chidambaram. Tamilnadu, India
  A Ayyasamy, Department of CSE, FEAT, Annamalai university, Chidambaram, Tamilnadu, India
  K Suresh Joseph, Department of Computer Science, Pondicherry University, Puducherry. India.

IoT is a set of physical and virtual items which are associated together by computer network for contact and sensing or interconnected with internal and external surroundings [6] If we identify the IoT **r**elates to supply chain system, it is position to place of physical objects which are coupled together for tracking, tracing, sensing and motivating within stimulated background for efficient communication and monitoring the supply chain. Our goal is to apply IoT in Food Supply Chain Management for creating communication between supply chain elements and procedures, classifying items and products automatically, tracking stream of products at every stage, presenting whole information during the entire supply chain process and achieving precision to overcome provocation of conventional supply chain [7] . RFID, Sensor, Barcode and wireless network technology have the potential to enhance the uneasiness of an IoT and numerous related parameters. In food traceability, the ecological situations are estimated using the sensors with food safety techniques [8] based on efficient methodology and rapid communication with the structure. Fig. 1 illustrates the applications of IoT in the recent years; the IoT has the interconnectivity with all the day to day life.

The connectivity has been done within the Transport systems, Agriculture sector, Energy utilization, Security and Privacy, Management of Building, Embedded systems, Wire- less systems, Pervasive computing, Wireless Sensor network, Smart cities and Healthcare applications [9] . Conventionally, perfect information is not often available in FSCM due to difficulty and expertise restrictions, usually only a segment of food supply chain issue can be tracked yet with incomplete tracing data [10] .

RFID has the facility to follow and sketch products in syn- chronized for the protection and quality of food. Nevertheless, most of the RFID executed very fast when compared to other insecurity limitation. In this paper assessment is prepared to determine out which type of decision is the best for Bayes classifiers algorithm. In this option there are 4 types of re- striction like complete test set, preparation set, proportion spilt and cross substantiation. The cross substantiation parameter is to compute the data set values. We progress with the food defect difficulty that may be cause by various reasons, such as bacterial fault, chemical fault, etc [12] .

The major contribution of the paper is

To build the quality based methodologies for food supply chain organization process from the producer to the con- sumers, assisting and revising whenever any irregular food condition.

The manufacture process is maintained by current sufficient data and guarantee to all the consumers.

Influencing these data to obtain an entire representation over contamination conditions in the complete network, such as the contamination basis and the other involved foods that require to be recalled.

The rest of this research is prepared as follows: A Literature Survey about internet of things and its application in FSCM presented in Section 2. Section 3 illustrates basic concept of IoT and FSCM and framework of Supply Chain Traceability and also comparison of Bayes classifier algorithm is proposed. Experimental results are also presented in Section 4. Conclusion and the future direction of the research presented in Section 5.
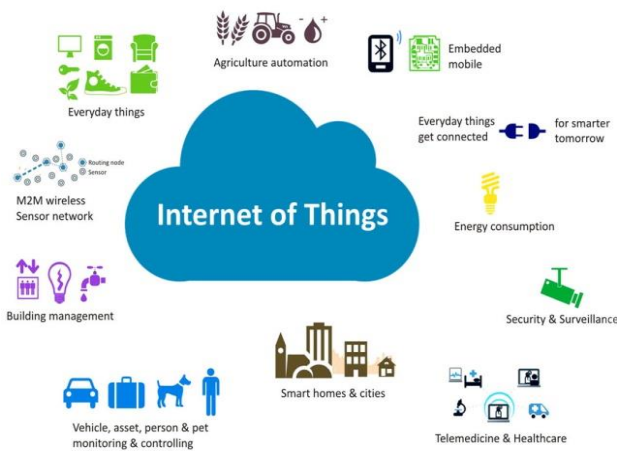

Fig. 1 IoT applications

II. LITERATURE SURVEY

A review on different application of IoT and its purpose in supply chain processes presented in this section. To group the most relevant literature to our research we have investigated Google scholar and also searched some publisher websites such as Springer, Elsevier, IEEE, and Emerald etc. The influence of Internet of Things (IoT) on various procedures of supply chain traceability is unknown. Since the entire food supply chain is divided into 5 associates: Farm; Processing; Warehousing; Distributor; Retail, then we will demonstrate the influence of IoT on each stage with feature in our literature review.

Every condition of the food supply chain traceability system has been conceded and demonstrated carefully to progress the safety of the food. HACCP is a preventive methodology [13] to remove the chemical contents in the production system. The smart background [14] developed which consist of transmitting the data onto the elegant network of Internet of Thing (IoT). For success the right decision, decision making model implemented on the data gathered from IoT devices by the business logic. Conclude that the data diagnostic in a business field provides the right decision at the right time. Moreover, it is the booming key in business [15].

Diverse technologies and data model have been experimentation in the various industries. In the food industry, it is seems to quite difficult to do so. It recommended a realistic structure that captures food giving out from the farmer to towards person major cuts in the

single entry in the framework [16]. A modern data with barcode and RFID is established in the part of decision support system for taking decision in correct time. In this system, several types of sensors are operated to confine the real-time data on temperature, humidity, and light. The data is accumulated and can be additionally analyzed how RFID can be implemented to outline the items in the FSCM [17].

Supply chain traceability in food manufacturing industry can gathering from division of the manufacture to the complete supply chain that includes the basic material provide to the ultimate utilization market [18]. A decision support system (DSS) that can check food worth, forecast temperature and humidity propose and execute improvement strategies in close to real-time and then re-examine situations to make a decision if additional action is needed such as incremental monitoring and employing enhanced strategies [19]. IoT is used in numerous real-time applications using Machine-to-Machine transmission [20], health support systems [21], Smart home based applications [22] and Industry automation system [23]. During the last decades, customer belief in the food manufacturing was very much damaged because of preserving the nutritional attribute of products during transportation from producer to customer [24]. There are lots of developed utility-driven industrialized systems with diverse IoT applications of sensors to supervise abnormal conditions [25]. With the rising popularity of sensors, barcode, actuators, and RFID tags, IoT applications produce enormous amounts of rich data per day [26, 27]. In few industrial provinces, food manufacturing plans can be considered by the exposed data and knowledge, which can be utilized to raise and expand maximum utilities [28]. Several analytic technologies are used in lots of diverse domains and they offer influential ways of discovering helpful, significant, and embedded data from particularly big datasets [29].

As a result, additional increasing consumer anxiety over the safety and quality of food have stressed more consideration from logical and industrial areas[30]. In reaction to increasing food security problems, many IoT technologies, such as RFID, sensor, barcode and various network technologies are useful to supply chain traceability and monitoring the system [31]. Most recent years an increased interest in the IoT, smart connected things in directive to implement the different outlook of the supplier, manufacturer and customers and conclude a standard which is not only supports performance and quality of the product also estimate the benefit of smart connected things [32].

IoT determine its function in all the fields, as recent structure of statement linking the various classification and procedure. IoT also called the internet of everything is extensive technology which is been out looked as a worldwide system of machines and devices able of communicating with each other. In increasing aspect and demands of modern cities, the association of food supply chain has turn into enormous difficult exceptional for quality of foods using the IoT based food traceability [33]. Complementary concerns also happen from this dilemma regarding loss of quality accuracy due to time and performance of tracing scheme.

Barcode and sensor-based IoT data is being practical to food chains from their incline of foundation to their measures idea in processing plants, storage warehouses, allocation points, and in the provisions. The laterally tracking and traceability that sensors and barcodes supply facilitate store chains, food brands, and goods deliver networks to rapidly recognize points of derivation and allotment if it's exposed that food is contaminated [34].

Bayes Classifiers Algorithm utilized experimental results for numerous descriptions of the Naive Bayes classifiers and a way of improving it using nearby subjective learning. And in the end, it shows how the management of Multinomial Naïve Bayes can be improved using locally subjective learning [35]. The rapid improvement of IoT and significant technologies has completed it more possible and practical to perform a multifaceted numerical assignment produced by a Bayesian Network representation [36].

Provision and farming products from manufacture to consumer of the entire process engages the manufacture, processing, stuffing, shipping, cargo space, shelf display and utilization, every connection is probable to carry the lacking confidence reasons [37]. Now-a-days approximately everybody is accomplishment unfair by the foodstuff that they are all eat, it is not concerning useless items and packing foods, the vegetables we are in position to suggest measure since they affected by temperature, moisture and circumstance throughout supply chain. Most of customers only give concentration to the data offered on the packing i.e. the quantity of constituents used and their nutritional value [38] and ignoring the ecological environment to which these packet are issued. In direct to manage the food before accomplishment fully to consumer, using IoT which permits tracking, tracing and managing the entire process of food supply chain such as supplier, exporter and consumers.

III. PROPOSED METHODOLOGY

Every food traceability management techniques projected in related work has own compensation as per individual process. Each and every one considered to gather individual customer requirements in diverse user fulfilment issues. The proposed system is to construct the food quality based technique for traceability process from the producer to the consumers, supporting and revising whenever any unbalanced food condition, save manual efforts and enhance high quality of food.

3.1 Food Supply Chain Work design
A typical food supply chain work design is offered in Fig.2 from the producer to the characteristics of the customer and the common description for the food traceability process. The traceability is point out that the farmers associated with the manufacturers and the transport parameter plays a significant role to create the food supply chain traceability.
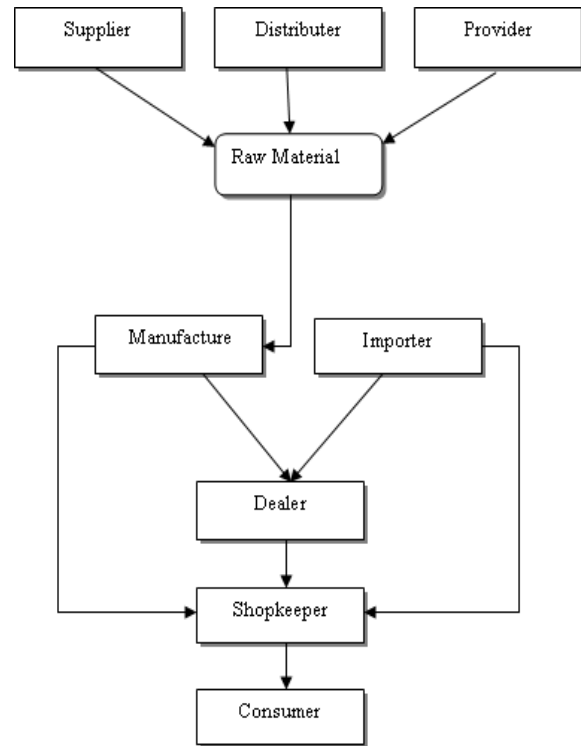

Fig. 2 Food supply chain work design

In the producer connection, initial of all, setting surroundings of farming always be calculated, as well as the excellence of the soil, water, air, etc. All realistic method ought to exist engaged to assurance the security of the rising and connected checking data must be required. In the dispensation connections, processing is feasible to take place at a dissimilar site. Consequently processing background and tools assessment will be compulsory. All resultant processing activities should be established according to best quality practices. In addition, using preservatives and possessions should be suitable for their designed rationale. In the Warehousing, supply chain tools have to be accurately sustained. Additionally, each and every one warehouse preparation like recording background of the storeroom, excellence and cargo space time of goods should be based on best quality practices. In the delivery correlation, all frozen tools, like truck, should be accurately sealed as well. Finally, in the retail phase, all retail administration practices like via the refrigeration; assessment the freshness period; alternative expired goods should be replaced according to best working practices and guarantee as an everyday manuscript.

3.2 Abstract structure of Food Supply Chain Traceability Design
Traceability engages the capacity to classify at any particular step of the food supply chain (from production to release) from one step back and one step forward. Traceability processes guarantee that foods are traced and tracked entire the supply chain management. Traceability is essentially significant for food safety as well as quality aspects. Traceability refers to the procedure that traces the work flow of foods throughout the manufacture, processing and delivery stages.
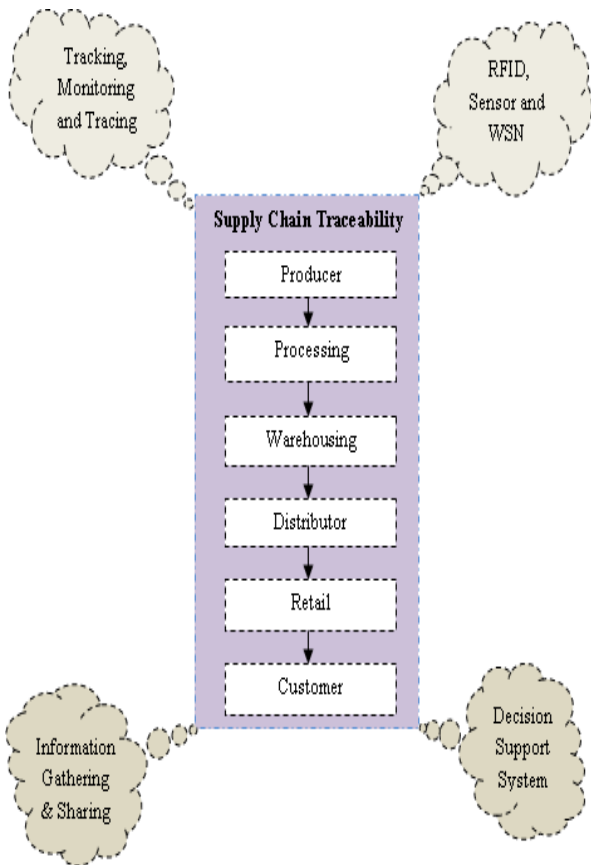
Fig. 3 Conceptual Framework of Food Supply Chain Traceability System

Now it is point to suggest a conceptual framework of food supply chain traceability system that distinguishes the tracing requirements in different stages: Tracking, monitoring and tracing, RFID, sensor, barcode and wireless network technology, Information gathering and sharing and finally Decision support systems. At the trailing, checking and tracing, unique trailing data of the objective flow of a central issue is collected using RFID, sensor, barcode and wireless network technology and group of information should be sharing across the network to take right decision for the quality of the product based on the decision support system.

3.3 Intellectual Transition Architecture of Food Process Traceability

Figure 4 demonstrates the proposed Intellectual transition architecture of food process traceability into the Decision Support System (DSS). In this representation, directly food manufacture data are composed from the sensor on the network system which streams of trailing items, machineries, cargoes etc. This fundamental data is implemented to create the correct decision for method traceability, along with the skilled information and the various databases, guided by the decision support system. Minimally call this network system as process-based traceability network based on discover the most possible tracking outputs.
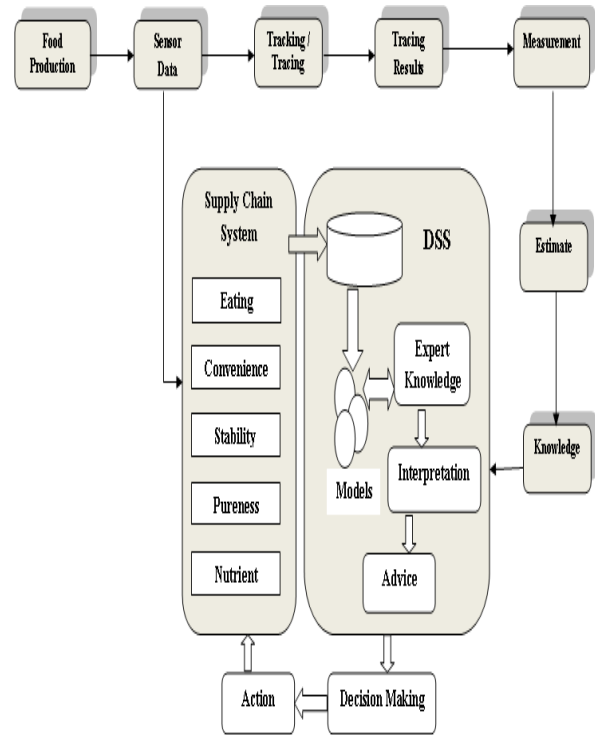

Fig. 4 Intellectual transition architecture of food process traceability

The above outputs are calculated and estimated to ensure for legality and will be operated to produced the facts that is then accumulated at the limited database. This database is then encouraged to development the traceability model. DSS join together all suitable data near the tracking difficulty, jointly with the sensor data; human knowledge and decision support in order to association all the decision interaction that exist in the production procedure.

The DSS supply the information and control to the user but the user creates the decisions pattern is method to construct a 'decision support' as a substitute of a decision. DSS assists the decision creator chooses among existing actions as long as additional information. A decision outputs in an action to be performed within the supply chain environment. Following the action is employed; the surrounding is again observed to start a novel cycle of information flow. Now it is time to discuss the improved Bayes Classifiers algorithm based process traceability model.

3.4 Bayes Classifiers algorithm

By the help of Bayes classifier technique we investigate the most excellent algorithm for the Food Supply chain dataset based on the cross substantiation attributes. Cross substantiation is the procedure of training learners via one part of data and trying it using a diverse set. Attributes tuning is the procedure to choosing the standards for a model's parameters that exploit the precision of the model. The work flow diagram for the relative investigation is shown in Figure 5.
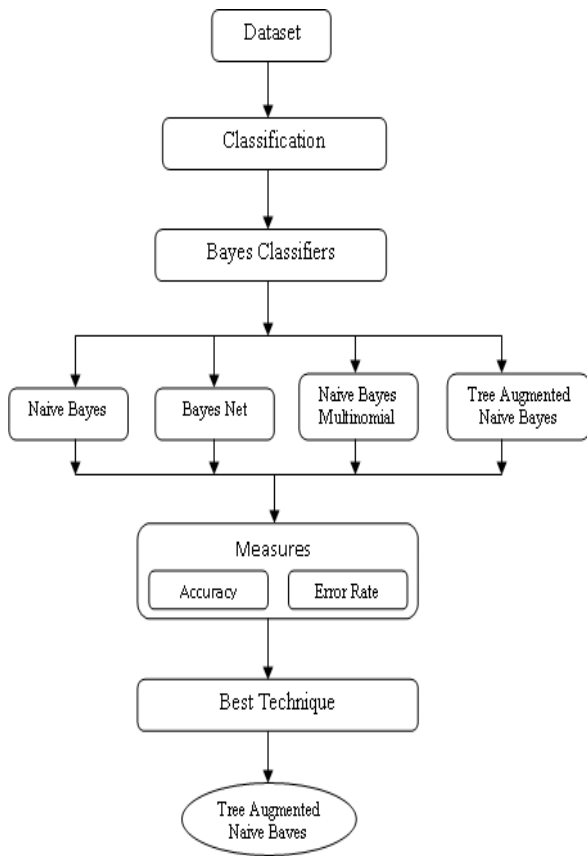
Fig. 5 Work flow diagram

3.4.1 Dataset

Food supply chain management plays a significant responsibility in our day by day since it provisions us with the essential for our life .However, uneconomical and unsuitable management systems may cause huge number of food losses and food safety. The Food supply chain dataset has been composed with help of the data world. This dataset contains 8791 instance and 50 attributes. The weka tool is executed for explore the management of the Bayes classifier algorithm.

3.4.2 Classification

The categorization processes cluster the data into the module on the foundation of their difference. A figure of the categorization techniques are Decision tree Classifier, Neural Network Classifier, Naïve Bayes Classifier, Decision forest and so on. Everyone can use technique make exploit of the knowledge algorithm to create the depiction that most excellent hysterics the connection involving the predictors and the estimate. In this paper the Bayes Classifiers algorithms are appraise to anticipate the algorithm is mainly suitable for the Food Supply chain dataset. In the Bayes classification technique, the four algorithms are compared namely Naive Bayes, Bayes Net, Naive Bayes Multinomial and Tree augmented Naive Bayes to determine out which is fits efficiently for the Food Supply chain dataset.

3.4.3 Bayes classifiers

Bayesian classifiers are variety of probabilistic graphical depiction that can be used to build demonstration from data and/or expert view. They can be implemented for a broad range of tasks including forecast, anomaly recognition, diagnostics, automated nearby, logic, time series calculation and decision making under ambiguity. The Bayes classification algorithms are used to find the most excellent algorithm for the Food Safety concerns and they are as

follows.
1. Naive Bayes
2. Bayes Net
3. Naive Bayes Multinomial
4. Tree augmented Naive Bayes

3.4.3.1 Naive Bayes

Naive Bayes classifiers are collection of various classifier algorithms formulated on Bayes Theorem. It is based on a multiple algorithm however an ancestor's component of algorithms in which all contribute to a universal principle, i.e. each couple of characteristics being confidential is self-governing of depending each other. The Naive Bayes classifiers are extremely reliable, concerning a numerous of attributes linear in the numeral of constraints in a knowledge difficulty. The process of approximation the parameters of a distribution by maximizing a likelihood function, so that under the believed statistical model the experimental data is mainly possible preparation can be finished by compute a closed-form term, which takes normal time, somewhat than by costly.
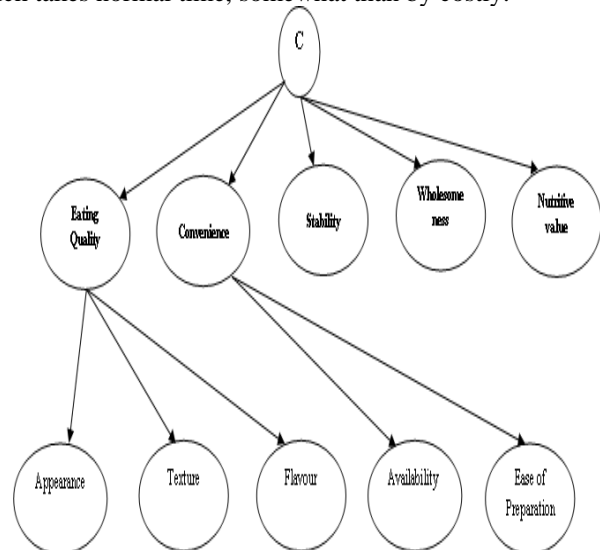

Fig. 6 Correlation between the attributes in Naive Bayes
With above figure to our idea, this plan can be unspoken as:

- We suppose that single attribute is reliant. For example, the Eating quality is individual 'Flavour' has not whatever thing to done with the Convenience being 'Availability or the wholesomeness being 'Purity' has denial effect on the Stability. Hence, the Attributes are supposed to be self-determining.
- Next, every attribute is given the identical influence (or consequence). For example, meaningful only Wholesomeness and Stability only can't anticipate the output accurateness. No attribute is unsuitable and whispered to be causally equal to the result. The postulation ended by Naive Bayes is rarely exact in current situations. In actual fact, the self-determination proposition is partially wrong but frequently facility well in preparation.

Bayes' Theorem discovers the possibility of an event occurring given the likelihood of one more event that has already happened. Bayes' theorem is known precisely as the following equation:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \ldots \ldots \ldots \ldots \ldots \ldots (1)$$

Where A and B are possible events. Essentially, it is difficult to locate probability of event A; given the event B is true. Event B is also expressed as facts. P(A) is the priori of A (the prior prospect, i.e. Probability of event before confirmation). The confirmation is an characteristic value of an unidentified case. P(A|B) is a posteriori chance of B, i.e. prospect of event after confirmation is proved. Now, with consider our dataset, we can pertain Bayes' theorem in following way:

$$P(b|A) = \frac{P(A|b)P(b)}{P(A)} \ldots \ldots \ldots \ldots \ldots \ldots \ldots (2)$$

where, b is class variable and A is a dependent attribute vector (of size *n*) where:

A= (a₁, a₂, a₃, a₄..........., aₙ)

Algorithm 1: Naive Bayes Construction procedure

Input: *data input a, b*

Output: *class A*

Parameters: *P(b) - class probability*

        P(a | b) - conditional probability

Procedure

Begin

    Read fictional training data

    For *every class*

        For *every attributes*

        Calculate P (b/a₁, a₂, a₃..... aₙ) α P(b)

$\prod_{i=1}^{n} P(ai|b)$

        Calculate P aᵢ | bᵢ)

        Calculate P (a | b)

        End

    End

        For *every data items*

        Calculate Posterior Probability

        Calculate class X using P(a) + P(b) = 1

        End

Return

    Class

End

At the instant, it's special occasion to put a naive postulation to the Bayes' theorem, which is, self-determination among the characteristic. So now, we split verification into the self-determining parts.

Now, if any two events A and B are independent, then

$$P(A, B) = P(A) P(B) \ldots \ldots \ldots \ldots (3)$$

Hence, we accomplish to the result:

$$P(b|((a1, a2\ldots, an) = \frac{P(a1|b)P(a2|b) \ldots \ldots P(an|b)P(b)}{P(a1)P(a2) \ldots P(an)} \ldots (4)$$

this can be articulated as:

$$P(b|((a1, a2\ldots, an) = \frac{P(b) \prod_{i=1}^{n} P(ai|b)}{P(a1)P(a2) \ldots P(an)} \ldots \ldots (5)$$

Now, as the denominator residue even for a known input, we can take away that term:

$$P(b|((a_1, a_2\ldots, a_n) \propto P(b) \prod_{i=1}^{n} P(ai|b) \ldots \ldots (6)$$

Now, we require building a model to locate the possibility of well-known place of contribution for all probable ideals of the class unpredictable *b* and decide up the production with highest prospect. This can be articulated exactly as:

$$b = argmax_b P(b)$$
$$\prod_{i=1}^{n} P(ai|b) \ldots \ldots \ldots \ldots \ldots \ldots (7)$$

So, ultimately, the assignment of manipulative P(b) and P(ai | b) where P(b) is also called class likelihood and P(ai | b) is called qualified possibility. The dissimilar naive Bayes classifiers disagree principally by the hypothesis they make regarding the allotment of P(ai | b).

Fig. 7 Flowchart for Conditional probability calculation

The forecast are completely based on particular data from confidence the extra data is obtained in which it enhanced. A new benefit is that Bayesian models are self- actuating, implication that when adjust the data, so we get the output. One of the extremely useful Bayesian is the Naive Bayes Classifier. It is purely based on the hypothesis; Bayesian theorem is predominantly well-matched when the dimensionality of the inputs is important.

3.4.3.2 BayesNet

Bayes Net can be implemented by using the Bayes theorem. Bayesian networks are probabilistic since they are created from possibility distributions and also develop the laws of possibility for forecast and irregularity detection and for way of thinking and diagnostics, conclusion making under insecurity and time sequence forecast. A Bayesian network describes the causal probabilistic association among a position of casual variables, their provisional dependences, and it near a concrete depiction of a joint possibility distribution. To compile up a Bayesian network originally conditional probability of every node must be measured. The acyclic graphs are exploited to differentiate the network. Previous than structure the network, it is understood that there are no misplaced ideals and all attribute values are hypothetical.

Algorithm 2: BayesNet Construction procedure

1. F ⟵ Ø
2. P ⟵ Probability Tables (F, C)
3. A ⟵ (u, F, P)
4. Attain ⟵ ∞
5. do:
    (a) Max attain   Attain
    (b) for each attribute pair (i, j) do
    (c) for each F′ ∈ {F U {I ⟶ J},
              ⟶ F - {I   J},
          F - {I ⟶ J} U {J ⟶ I}}
    (d) T' ⟵ Probability Tables (F', C)
    (e) A' ⟵ (u, F ', T')
    (f) newattain ⟵ AICattain(A', C)
    (g) if newattain > Attain
          then A ⟵ A'
          Attain ⟵ newattain
6. While Attain > Maxattain
7. Return A

3.4.3.3 Multinomial Naive Bayes

The Multinomial event illustration indicated to as Multinomial Naive Bayes usually performs well the multivariate one and it also start to appraise usefully with more devoted event models. For assume the text categorization dilemma, Naive Bayes Multinomial text algorithm is used.

Algorithm 3: Naive Bayes Multinomial Construction procedure

NBMULTINOMIAL (A, B)
1. E ⟵ EXPRESSION (A, B)
2. N ⟵ COUNT (D)
3. For each c ∈ A
4. Do NC ⟵ COUNTINCLASS (B,c)
5. Prior1[c] ⟵ Nc / No
6. text$_c$ ⟵ CONCATENATETEXTOF ALLDOCSINCLASS(B,A)
7. for each t ∈ E
8. Do T$_{ct}$ ⟵ COUNTTOKENOFTREM (text$_c$, t)
9. Do condprob[t][c] ⟵ T$_{ct+1}$ / $\sum \sum t'$ (Tct'+1)
10. return v, prior1, condprob1
        APPLY NBMULTINOMIAL   (C, E, prior1, condprob1, d)
        1. w ⟵ EXTRACTTOKENSFROMDOC(V,B)
        2. for each c ∈ A
        3. do attain[c]   log prior[c]
        4. for each t ∈ w
        5. do attain [c]+ = log condprob1[t][c]
        6. return argmax c ∈ A attain[c]

3.4.3.4 Tree Augmented Naive Bayes

In the Naive Bayes representation the statement is that all the attributes are self-determining of each other in the class. In certainty, the attributes are infrequently self-regulating of each other. In most number of cases attributes are straight or not directly dependent on each other. Even though this naive assumption not often holds up, the Bayesian Network model may still execute moderately well. If we take into version some of the well-built conditional dependences between the attributes, the presentation of Naive Bayes classification can be improved. A joint probability distribution is exactly characterized by a Bayesian Networks, but a complete Bayesian Networks is expensive to construct and costly for assumption.

In a Bayesian Networks we confine the inter relationship between the random variables as well as the self-determination between them. Using this we can compute the conditional probabilities of each random variable given its parents. Each random variable is independent of its non-descendants given its parents. For a improved classification presentation, we require an expansion of this model that also comprises some dependencies between the random variables, but not constructing a complete Bayesian Networks. An enhanced classification representation can be accomplished by augmenting the Naive Bayes model.

An augmented Naïve Bayes model keeps the fundamental construction of the model but augments it by count the edges among the attributes to take sketch of the information of relationship between the attributes. This procedure will enhance the compute complexity of the system but in the same time classification will more accurate. One such model is the Tree-Augmented Naïve Bayes (TANB) Model; we limit the level of communication between the random variables and the essential configuration of Naive Bayes representation is taken. There are straight edges between the class node and all the attributes. Consequently, it will take into concern all the attributes while manipulating the $P(C|A_1,A_2…..A_N)$.

In addition of that, each random variable is associated with other random variable throughout a direct edge apart from dedicated attribute called the root. Except for the class, each variable in the network will have one or more parents: one is class node and the other one is another random variable. As communication between the attributes have been inadequate to the calculate complexity of this representation is very much concentrated. Thus, TANB representation sustains the strength and computational complexity of the Naive Bayes model and at the similar time shows improved accuracy.

3.4.3.4.1 Structure of a TANB model for Food supply chain evaluation

TANB enforces the constraint that the numeral of correlation connecting the attributes associated with the classification is restricted to single. Below figure shows an example TANB model for the Food supply chain evaluation data set. In this Food Supply Chain evaluation dataset which try to estimate whether a food is suitable specified attributes associated to quality, safety and price. For this dataset if we consider the attribute like stability of the food, it is closely related to the appearance of the food. If the appearance of the food is good, then the food is generally safer and similarly if the appearance is not bad generally the safety features related to the food is lesser. In case of Naive Bayes representation, these would be measured as two separate independent events, so it will over castigate the class label. But in the case of TANB model the significant correlation between these events.

$$FI(I;J) = \Sigma_{i,j} P(i, j) * (\log (P(i, j)/P(i)P(j))) …………..(8)$$

If the two variables are given the well-known information will examine a lot of one variable provide information about the other.
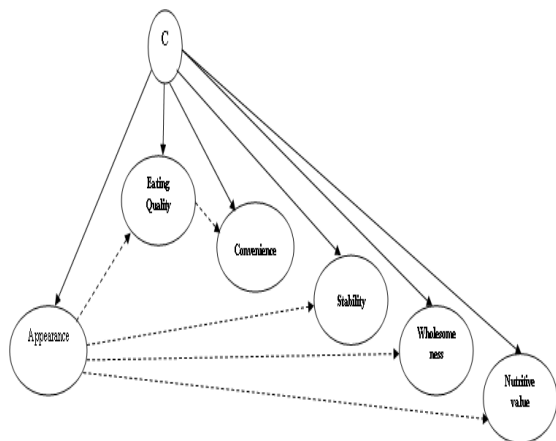
Fig. 8 Correlation between the attributes in TANB

Note down that the above figure, the edges connecting the class labels and attributes are characterized using solid lines and edges connecting the attributes are characterized using the dotted lines. Here all the attributes apart from the appearance attribute has two parents, so it is measured as the root. If we take away all the edges from the class label to the attributes a tree structure can be pictured. All the edges can be observed direct away from the root.

3.4.3.4.2 Implementation

In order to discover the significant correlated attributes within this representation familiar information is utilized. The familiar information is considered between each pair of attributes which structure the influence of the edges. The edges are supplementary between the attributes which are extremely mutually dependent. If there are M attributes in the organization, there will M nodes in the tree structure and there will be M-1 edges essential to attach all the nodes within the diagram. The sum of the familiar information on the edges within the diagram should outline maximum weight spanning tree. The familiar information between two random variables I and J is defined by the following

$$FI(I;J) = \Sigma_{i,j} P(i, j) * (\log (P(i, j)/P(i)P(j)))...............(8a)$$

If two variables are given the familiar information will analyse a lot of one variable supply information about the other. Conditional familiar information given the class label is utilized to build the tree structure in TANB model. The conditional familiar information used to construct the tree structure is given below:

$$FI(I;J \mid K) = \Sigma_{i,j,k} P(i, j, k) * \log (P(i, j \mid k)/P(i \mid k)P(j \mid k))......(9)$$

The tree can be constructed for the TANB model by using the following algorithm:

Algorithm4: TANB Construction procedure

Step 1:
Compute $FI_p(A_x, A_y \mid C)$ for each pair of attributes where $x \neq y$.

Step 2:
An undirected graph is constructed with nodes from A1,A2, ….,AN. Allocate the weight of edge between A_x, A_y using $FI_p(A_x, A_y \mid C)$.

Step 3:
Construct the highest weighted spanning tree.

Step 4:
The undirected graph is renewed into a directed graph by deciding a random variable as root and straight all the edges outward from it.

For the period of classification the posterior probability of every one of the class label P $(C|A_1A_2.....A_N)$ is calculated. The class label with highest posterior probability is used to categorize the test record. As discussed formerly, we employ Laplace correction to moderate the condition where given class and characteristic does not emerge jointly in the training dataset. The procedure for TANB classification is given by:

$$P (C \mid A_1, A_2… A_N) = P(C) \cdot P(A_{root} \mid C) \Pi_i . P(A_i \mid C , A_{parent}).......(10)$$
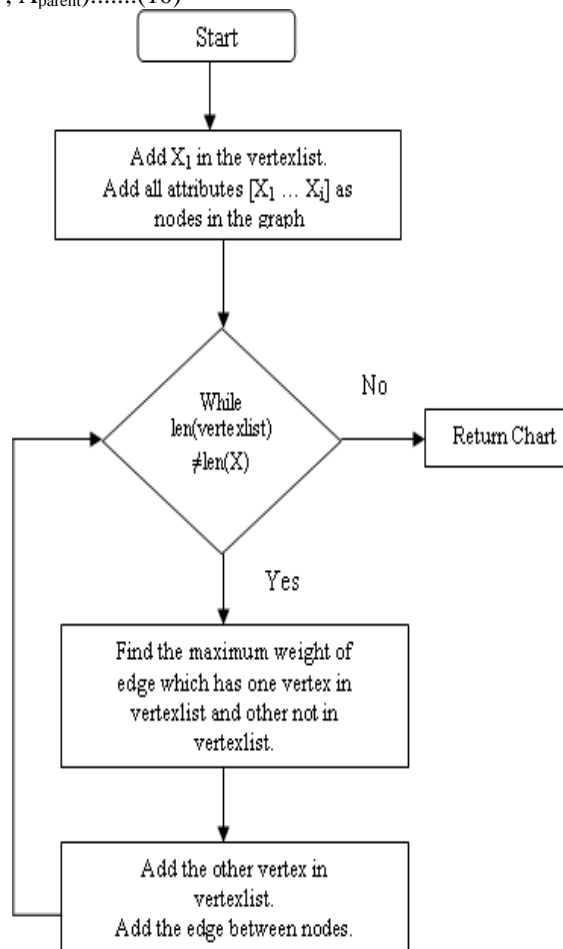

Fig. 9 Flowchart for Constructing Tree in TANB model

To build a TANB representation, the pair wise familiar information is considered between all nodes. The highest weighted tree is then constructed using TANB algorithm. Also, the previous probabilities and conditional probabilities want to be designed, which is very comparable to Naive Bayes representation. Then the categorization of the experiment records is done.

IV. EXPERIMENTAL OUTPUTS

Based on the above investigational process is considered by means of the performance concerns such as the classification correctness and fault rates. And also calculate the qualified examination for the Food supply chain dataset to forecast the optimum algorithm. The accurateness appraises and the performance concerns by group for the Bayes classifiers are represent in Table 1.

Table 1: Evaluation of performance factors for Bayes classifiers algorithms

Sl.no	Algorithm	Correctly classified instances (% value)	Incorrectly classified Instances (% value)
1	Naive Bayes Classifier	82.6037 %	17.3963%
2	Bayes Net Classifier	81.3333%	18.6667%
3	Naive Bayes Multinomial Classifier	78.7037%	21.2963%
4	Tree Augmented Naive Bayes (TANB) Classifier	84.2087 %	15.7913%

Based on Table 2, it is incidental that for Tree Augmented Naive Bayes (TANB) on cross justification attributes, the F-Measure, TP rate, ROC curve, Precision and the Kappa values are greater than the other three algorithms such as the Naive Bayes, Bayes Net and Naïve Bayes Multinomial. The evaluation of performance process for Bayes classifier has shown in Figure 11 and the accuracy measure for the Bayes classifiers is shown in Table 2.

Table 2: Comparison of accuracy measure for Bayes classifiers algorithms

Sl.no	Algorithm	TP Rate	Precision	F-Measure	ROC Curve	Kappa value
1	Naive Bayes Classifier	0.826	0.83	0.826	0.836	0.6459
2	Bayes Net Classifier	0.813	0.81	0.813	0.813	0.6312
3	Naive Bayes Multinomial Classifier	0.787	0.79	0.787	0.797	0.5267
4	Tree Augmented Naive Bayes (TANB) Classifier	0.842	0.84	0.842	0.862	0.6812

Based on Table 2, it is conditional that the Tree Augmented Naive Bayes (TANB) algorithm has advanced classification accuracy than the former classification algorithms such as the Naive Bayes, Bayes Net and Naïve Bayes Multinomial. The comparison of the accuracy measures for the Bayes classifiers is shown in Figure 10 and the error rate measures for the Bayes classifiers are shown in Table 3.
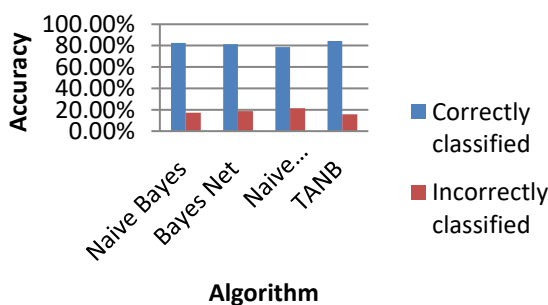
Accuracy Measure

Figure 10: Comparison of Accuracy Measure for Classification algorithms

For correctly classified instances, it is conclude that TANB performs 1.6 % better than Naive Bayes algorithm, 2.8% better than BayesNet and 5.5% better than Naïve Bayes Multinomial. Similarly for incorrectly classified instances it

is indirect that TANB performs 9.2% better than Naïve Bayes algorithm performs, 15.4% better than BayesNet and 20.5% better than Naïve Bayes Multinomial.

For TP rate, it is indirect that TANB performs 1.6% healthier than Naive Bayes algorithm, 2.9% healthier than BayesNet and 5.5% healthier than Naïve Bayes Multinomial. For precision it is indirect that TANB performs 1% healthier than Naive Bayes algorithm, 3% healthier than BayesNet and 5% healthier than Naïve Bayes Multinomial.
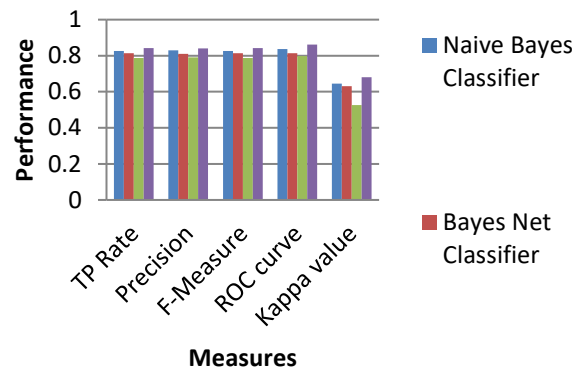
Performance Measures

Figure 11: Evaluation of performance factors for Bayes classifiers algorithms

For F-measure it is indirect that TANB performs 1.6% healthier than Naive Bayes algorithm, 2.9% healthier than BayesNet and 5.5% healthier than Naïve Bayes Multinomial. For ROC Curve it is conclude that TANB performs 2.6% healthier than Naive Bayes algorithm, 4.9% healthier than BayesNet and 6.5% healthier than Naïve Bayes Multinomial. For kappa Value it is indirect that TANB performs 3.53% healthier than Naive Bayes algorithm, 5.01% healthier than BayesNet and 15.4% better than Naïve Bayes Multinomial.

Based on Table 3, it is conditional that the TANB algorithm has the smallest error rates than the Naïve Bayes classification, Bayes Net and Naïve Bayes Multinomial. The assessment of the error process for the Bayes classifiers is shown in Figures 12 and 13.

Table 3: Evaluation of error rate measures for Bayes classifiers algorithms

Sl.no	Algorithm	MAE	RMSE	RAE	RRSE
1	Naive Bayes Classifier	0.1863	0.3607	37.7196	72.5867
2	Bayes Net Classifier	0.1947	0.3604	39.4261	72.5214
3	Naive Bayes Multinomial Classifier	0.2664	0.4869	53.9452	97.9937
4	Tree Augmented Naive Bayes (TANB) Classifier	0.1724	0.2891	35.2143	71.2581

For MAE, it is inferred that TANB performs 7.46% better than Naïve Bayes algorithm, 11.45% better than BayesNet and 35.28% healthier than Naïve Bayes Multinomial. For RMSE, it is incidental that TANB performs 19.85% healthier than that Naïve Bayes algorithm, 19.78% healthier than BayesNet and 40.62% healthier than Naïve Bayes Multinomial.
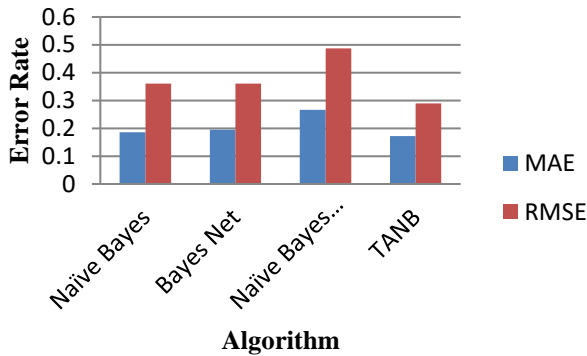
Error Rate Measures

Figure 12: Evaluation of Error rates for Bayes classifiers algorithms

For RAE it is conditional that TANB performs 6.64% healthier than that Naïve Bayes algorithm, 10.68% healthier than BayesNet and 34.72% healthier than Naïve Bayes Multinomial. For RRSE Curve it is conditional that TANB executes 1.83% healthier than Naive Bayes algorithm, 1.74% healthier than BayesNet and 27.28% healthier than Naive Bayes Multinomial.
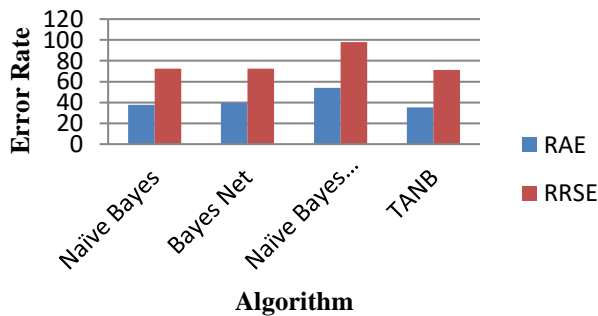
Error Rate Measures

Figure 13: Evaluation of Error rates for Bayes classifiers algorithms

The sensitivity of the Food Supply Chain management is experimentally evaluated that the proposed algorithm has the improved sensitivity according to the probability value. Finally, TANB is the most accurate classifier compared to all other classification algorithm based on performance, accuracy and error rate measures and Figure 14 shows the overall performance among the 4 algorithms.
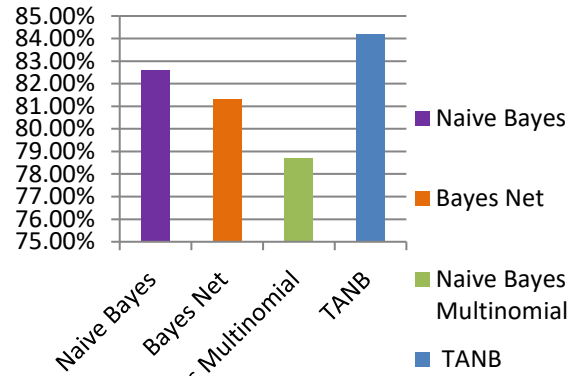
Overall Performance

Figure 14: Overall Performance

V. CONCLUSION

This paper examined the performance of 4 Bayes classifiers algorithms namely Naive Bayes, Bayes Net, Naive Bayes Multinomial and Tree augmented Naive Bayes. The Food safety datasets is used for manipulating the accomplishment by using cross validation constraint based on the class attribute. The algorithms are investigating based on the performance factors such as classification accurateness and fault rates. Based on the experimental results, it is basically Tree augmented Naive Bayes algorithm executes improved than other algorithms. The proposed algorithm ensures that best quality of food for the customer and also reduces the health risk of food contamination which leads to sudden death of human beings. Using IoT, the system is effectively monitoring and analyzing the quality of food products during the food supply chain traceability in order to provide high intensity of food safety to all.

Future Enhancement

In the future enhancement, the Bayes classification algorithm can be researched on other datasets to achieve more successful results. Also the Bayes classification algorithms can be analyzed by means of considerations such as the training set; percentage split, and supplied test set.

Ethical Responsibilities of Authors

The authors declare that they do not have any conflict of interests. This research does not involve any human or animal participation. All authors have checked and agreed the submission.

REFERENCES

1. Aung et al. , "Internet of Things: A review on current technologies, protocols, and applications," IEEE Commun. Surveys Tuts., vol. 17, no. 4, pp. 2347_2376, 4th Quart., 2015.
2. Atzori et al. "The internet of things: A survey," Computer Networks, vol. 54, no. 15, pp. 2787–2805, 2010.
3. Wen et al. "Design implementation and evaluation of an Internet of Things (IoT) network system for restaurant food waste management", Waste Management, vol. 73, pp. 26-38, Mar. 2018.
4. Scholten et al. (2016), "Defining and analyzing traceability systems in food supply chains", in Espiñeira, M. and Santaclara, F.J. (Eds), Advances in Food Traceability Techniques and Technologies, Elsevier, New York, NY, pp. 9-33

5. Xiao, et al.., Research on a Food Supply Chain Traceability Management System Based on RFID. Journal of Agricultural Mechanization Research. 2012, 34(2), 181-184.
6. Abad, E., et al., RFID smart tag for traceability and cold chain monitoring of food: demonstration in an intercontinental fresh fish logistic chain. Journal of Food Engineering. 2009, 93(4), 394-399.
7. Y. Liu, et al. "Exploring data validity in transportation systems for smart cities," IEEE Communications Magazine, vol. 55, no. 5, pp. 26–33, 2017.
8. Abdulah, et al. (2011) 'Smart packaging: sensors for monitoring of food quality and safety', Sensory and Instrumentation for Food Quality, Vol. 5, No. 3, pp.137–146.
9. Folinas, et al.., Traceability data management for food chains. British Food Journal. 2006, 108(8), 622-633.
10. Mishra, PM. Internet of Things and Bayesian Networks. Available from http://www.analyticbridge.com/profiles/blogs/internet-of-things-andbayesian- networks. July 9, 2014
11. Chen et al. Research on agricultural products cold-chain logistics of various applications in International conference on computer and computing technologies in agriculture, China: Springer International Publishing, 2014, pp. 247–253.
12. Ryan, J. M. (2014) Guide to Food Safety and Quality during Transportation: Controls, Standards and Practices, Elsevier.
13. N. Mai, et al. "Benefits of traceability in fish supply chains—case studies," British Food Journal, vol. 112, no. 9, pp. 976–1002, 2010.
14. D. Zhang et al., "The development and standardization of testing methods for genetically modified organisms and their derived product," Journal of Integrative Plant Biology, vol. 53,no. 7, pp. 539–551, 2011.
15. S. Allata, et al. "Implementation of traceability and food safety systems (HACCP) under the ISO 22000:2005 standard in North Africa: The case study of an ice cream company in Algeria", Food Control, vol. 79, pp. 239-253, Sep. 2017
16. Z. Sheng et al., "A survey on the ietf protocol suite for the internet of things: Standards, challenges, and opportunities," IEEE Wireless Communications, vol. 20, no. 6, pp. 91–98, 2013.
17. Fenu, G. et al.. "RFID-based supply chain traceability system", 35th Annual Conference of IEEE, IEEE, New York, NY, pp. 2672-2677, 2009
18. Moe, T.: Perspectives on traceability in food manufacture. Trends Food Sci. Technol. 9(5), 211–214 (1998)
19. S. H. Sutar et al. ``Integration of smart phone and IoT for development of smart public transportation system,'' in Proc. Int. Conf. Internet Things Appl., Jan. 2016, pp. 73_78.
20. S. M. Riazul Islam et al. ``The Internet of Things for health care: A comprehensive survey,'' IEEE Access, vol. 3, pp. 678_708, 2015.
21. S. Feng, et al. ,'' IEEE Communication. Mag., vol. 55, no. 2, pp. 34_39, Feb. 2017.
22. I. E. Etim and J. Lota, ``Power control in cognitive radios, Internet-of-Things (IoT) for factories and industrial automation,'' in Proc. Annu. Conf. IEEE Ind. Electron. Soc., Oct. 2016, pp. 4701_4705.
23. A. Miles, et al. Browne (2018): IoT-based decision support system for monitoring and mitigating atmospheric pollution in smart cities, Journal of Decision Systems, DOI: 10.1080/12460125.2018.1468696
24. Kelepouris et al.: RFID-enabled traceability in the food supply chain. Ind. Manage. Data Syst. 107(2), 183–200 (2007)
25. M. M. Hossain, et al., "Towards an analysis of security issues, challenges, and open problems in the internet of things," in Services (SERVICES), 2015 IEEE World Congress on. IEEE, 2015, pp. 21–28.
26. L. Da Xu et al. "Internet of things in industries: A survey," IEEE Transactions on industrial informatics, vol. 10, no. 4, pp. 2233– 2243, 2014.
27. Meaghan Lee et al. , "Agricultural Production System based on IoT", 978-0-7695-5096-1/13, 2013 IEEE DOI 10.1109/CSE.2013.126.
28. Yan, B et al..: Development of traceability system of aquatic foods supply chain based on RFID and EPC internet of things. Trans. Chin. Soc. Agric. Eng. 29(15), 172–183 (2013)
29. S. Chen, et al. "A vision of iot: Applications, challenges, and opportunities with china perspective," IEEE Internet of Things journal, vol. 1, no. 4, pp. 349–359, 2014.
30. Jansen-Vullers et al..: Managing traceability information in manufacture. Int. J. Inf. Manage. 23(5), 395–413 (2003).
31. Banerjee et al. . A block chain future for internet of things security: A position paper. Digital Communication Networks, 4: 149-160.
32. Ray, P.P., 2016. A survey on internet of things architectures. J. King Saud Univ. Comput. Inf. Sci., 30: 219-319.
33. M. Trebar et al. "Towards RFID traceability systems of farmed fish supply chain," in Proceedings of the 19th International Conference on Software, Telecommunications and Computer Networks (SoftCOM'11), pp. 6–11, Hvar, Croatia, September 2011.
34. Jensen, F.V.: An Introduction to Bayesian Networks. Springer, New York (1996).
35. Kibriya, et al., "Multinomial Naïve Bayes for Text Categorization Revisited" http://hdl.handle.net/10289/1448
36. Food Safety Management System (2014) The Amber Valley, [Online]: http://www.ambervalley.gov.uk/health-and-social-care/food-safety/food-safety management-system.aspx.
37. Breathe Nikola, et al. , Internet of Things: a Review from "Farm to Fork", IEEE Conference June 2016.
38. Feng Tian. (2017). A supply chain traceability system for food safety based on HACCP, Blockchain & Internet of things. 2017 International Conference on Service Systems and Service Management. doi:10.1109/icsssm.2017.7996119.

AUTHORS PROFILE

S. Balamurugan received the MCA Degree from Bharathidasan University, Tiruchirapalli, Tamil Nadu and M.Tech. Degree in Computer Science and Engg. from PRIST University, Thanjavur, India. He is currently pursuing the Ph.D. in the Department of Computer Science and Engineering, Annamalai University, Chidambaram, India. His research interests include wireless sensor networks, Internet of Things (IoT), Web Technology and Client Server System. He is a member of professional bodies like ISTE.

Dr. A. Ayyasamy completed his B.E. & M.E. in Computer Science and Engineering from Annamalai University, Chidambaram, Tamilnadu, India in the year 2006 and 2008 respectively. He is working as Assistant Professor in Department of Computer Science and Engineering, Faculty of Engineering and Technology, Annamalai University from 2007 where he obtained his Doctorate in 2015. He has guided 20 under graduate students and 18 post graduate students.

Dr. K.Suresh Joseph is working as Associate Professor in Department of Computer Science in Pondicherry University. He is completed his B.E. & M.E. in Computer Science and Engineering from Bharathiar university and University of Madras respectively. His areas of interest are Soft Computing, NDN. Having more than 13 years of experience in Academic side and completed his doctorate in the year of 2013 in Anna University.