Building an IoT Framework for Connected Dairy

Anitha Ilapakurti
IoT & Data Analytics Group, Hanumayamma Innovations and Technologies Inc, Fremont, CA 94536
ailapakurti@hanuinnotech.com

Chandrasekar Vuppalapati
Computer Engineering Department
San Jose State University, San Jose, CA 95192
chandrasekar.vuppalapati@sjsu.edu

Abstract – Heat stress (HS) causes cows to produce less milk with the same nutritional input, which effectively increases farmers' production costs. The economic toll due to higher-temperature, heat stress is a $1 billion annual problem[i]. Not only in the United States, but also around the globe heat stress causes an adverse impact on dairy productivity. The opportunities, however, for the dairy industry is to electronically monitor cattle temperature and implement appropriate measures so that the impact of HS can be minimized. The U.S. Department of Agriculture estimates nearly $2.4 billion a year in losses from animal illnesses that lead to death can be prevented by electronically checking on cattle's' vital signs[ii]. This research paper recommends the most innovative electronic monitor framework, the 'Smart Connected Objects", aka, 'the Internet of Things (IoT)', that enables dairies to minimize the economic impact of HS and, at the same, capture the higher Return on Assets (ROA) & Return on Investment (ROI) by improving operational efficiencies. Happy Cow, more importantly, means happier, more profitable, dairy industry and richer and creamer dairy products. The proposed framework supports both offline and online dairy IoT. This paper presents a prototyping solution design as well as its application and certain experimental results.

Keywords— Heat Stress; Sensor Tag; IoT; CEP; Complex Event Processing; Internet Of Things; Dairy; BLE; Bluetooth Low Energy; iOS; Android; IoT reference architecture; streaming analytics; Decision Tree; Memory Data Cube; Machine Learning; Regression Analysis; Dairy Cattle; and e- commerce

I. INTRODUCTION

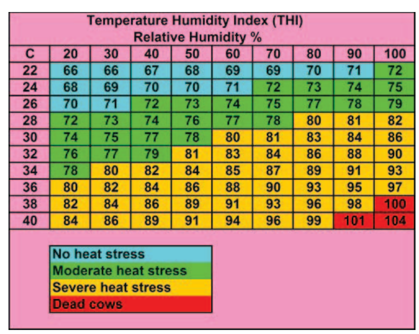
Dairy cows are homoeothermic animals and need to maintain a constant body temperature. They are sensitive to factors which influence their thermal exchange with the environment. These factors include air temperature, radiant temperature, air velocity, and relative humidity [1]. Air temperature and radiant temperature directly influence the heat exchange ability of the animal. Humidity can decrease heat exchange and have debilitating effects on the cow.

"Put simply, dairy cattle don't like the heat, and the stress that hotter weather places on them is close to a $1 billion annual problem", according to a national expert, Professor Robert Collier. He's a dairy specialist at the University of Arizona [2].

According to the findings of "The Effect of Climate Change on the Production Costs of the Dairy Industry in the United States" group project [3], "Heat stress causes cows to produce less milk with the same nutritional input, which effectively increases farmers' production costs. Furthermore, heat stress lowers the protein and fat content of milk. Higher temperatures additionally cause heat stress for dairy cows, leading to a reduction in milk yields. These impacts will pose additional burden on dairy farmers who operate on small profit margins." [3]

The comfort zone or thermos-neutral zone for a dairy cow is between 5°C and 25° C. 5° C is called the Lower Critical Temperature (LCT) and 25° C the Upper Critical Temperature (UCT). At temperatures below the LCT the cow will increase her dry matter intake to keep warm or convert feed to heat rather than produce milk. At temperatures above the UCT, cows have two main control strategies to maintain their thermal balance: (i) increasing heat dispersion and (ii). limiting heat production – both will reduce the milk productivity. [1]

According to research study[iii] conducted by Jeffrey and Richard, extension Dairy Specialists, University of Nebraska-Lincoln, a temperature of 100° F /37.7 ° C and 20 percent humidity is the range in which you begin "***serious measures"*** to ease the stress on the cattle. Some type of cooling should be started. The danger occurs as the temperature nears 100° F (37.7 ° C) and 50 percent humidity. The ***lethal range*** for cattle is 100V F and 80 percent humidity [4]. Finally, Legates *et al*., have demonstrated that significant breed differences were found for body temperature and respiratory rate [3].


Temperature Humidity Index (THI)
Relative Humidity %

C	20	30	40	50	60	70	80	90	100
22	66	66	67	68	69	69	70	71	72
24	68	69	70	70	71	72	73	74	75
26	70	71	72	73	74	75	77	78	79
28	72	73	74	76	77	78	80	81	82
30	74	75	77	78	80	81	83	84	86
32	76	77	79	81	83	84	86	88	90
34	78	80	82	84	85	87	89	91	93
36	80	82	84	86	88	90	93	95	97
38	82	84	86	89	91	93	96	98	100
40	84	86	89	91	94	96	99	101	104

No heat stress
Moderate heat stress
Severe heat stress
Dead cows

Figure 1. Temperature and Humidity Index [iv]

It is clear from the above studies that the lack of proper measures to ease the stress on the cattle and external environmental factors such as temperature and humidity can have an adverse effect on cattle's' internal vital signs and thus have a considerable impact on health and the productivity. Systematic monitoring of dairy cattle, on the other hand, would help in minimizing the stress effects. For instance, U.S.

978-1-4799-8128-1/15 $31.00 © 2015 IEEE
DOI 10.1109/BigDataService.2015.39

Department of Agriculture estimates nearly $2.4 billion a year in losses from animal illnesses that lead to death can be prevented with by electronically checking on cattle's heart rate, respiratory rate, digestion, temperature, and other vitals [6]. A New York Angels-backed pilot, importantly, is saving the agricultural industry a predicted $10 billion by monitoring the health of animals by a sensor-enabled smart pill. [7]

The Systematic electronic monitoring of dairy cattle can be achieved with the help of connected, albeit, Smart, internet enabled sensor objects. Gartner [8] defines the so called Smart Internet of connected objects as the "Internet of Things (IoT)". According to Gartner Research, IoT is "the network of physical objects that contain embedded technology to communicate and sense or interact with their internal states or the external environment."[8]

This paper addresses connected dairy based on IoT framework. Its major contribution is its proposed innovative solution that is based on Low Powered Bluetooth Framework, Embedded Sensors, Software Edge Analytics, and multi-dimension big data analytics. In this approach, we focus on Bluetooth low energy framework to connect to Sensors, retrieve the sensor data characteristics, apply edge processing that can filter, aggregate, enrich, and analyze a high throughput of data from the sensors to visualize health of dairy cattle in real time, detect urgent situations, and automate immediate actions.

The structure of this paper is as follows. Section 2 discusses the basic concepts and methods about IoT, Streaming, Bluetooth Low Energy Framework, and Edge Analytics. Section 3 presents our IoT analytics service system by focusing on its service framework. Section 4 discusses its related design and implementation decisions and Section 5 shows a case study. Conclusions and future work are summarized in Section 6.

II. UNDERSTANDING IOT FOR CONNECTED DAIRY

A. Internet of Things

1) Gartner Research [8] states "the Internet of Things (IoT) is the network of physical objects that contain embedded technology to communicate and sense or interact with their internal states or the external environment."

According to IBM [9], "the Internet of Things represents an evolution in which objects are capable of interacting with other objects. Hospitals can monitor and regulate pacemakers long distance, factories can automatically address production line issues and hotels can adjust temperature and lighting according to a guest's preferences, to name just a few examples."

As per an article written by to Porter and Heppelmann [10], "The increasing capabilities of smart, connected products not only reshape competition within industries but expand industry boundaries. This occurs as the basis of competition shifts from discrete products, to product systems consisting of closely related products, to systems of systems that link an array of product systems together. A tractor company, for example, may find itself competing in a broader farm automation industry."

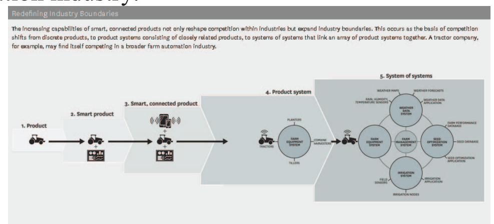


Figure 2: Source[v]

We can clearly see the "Dairy Industry of the future is more connected and data Centric. "

B. Bluetooth Low Energy Framework

Bluetooth low energy (Bluetooth LE, BLE, marketed as Bluetooth Smart[vii]) is a wireless personal area network technology designed and marketed by the Bluetooth Special Interest Group aimed at novel applications in the healthcare, fitness, beacons, security, and home entertainment industries. Compared to Classic Bluetooth, Bluetooth Smart is intended to provide considerably reduced power consumption and cost while maintaining a similar communication range [12].

C. Bluetooth Low Energy Sensor Tag

For our research purpose, we have used Texas Instruments' The Texas Bluetooth SensorTag [13]. The Bluetooth SensorTag demonstrates the low power capabilities of Bluetooth low energy (aka Bluetooth 4.0 and Bluetooth smart) [13].

Figure 3 TI Sensor Tag [ix]

The Bluetooth Sensor Tag provides real-time capture of Ambient Temperature, Object Temperature, Relative Humidity, Accelerometer, Relative Humidity, Barometer, and Compass. The Sensor tag fulfills the need to electronic monitoring of dairy cattle, Figure 4.

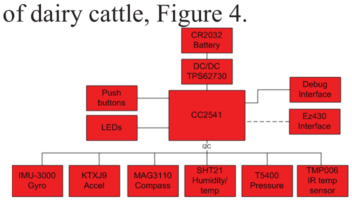


Figure 4: SensorTag Block Diagram[x]

D. Mobile Platforms and BLE Frameworks

Mobile operating systems, including iOS, Android, Windows Phone and BlackBerry, OS X, Linux, and

Windows 8, natively support Bluetooth Smart [14]. For Connected Dairy we support Android 4.3 or higher and iOS 7 or higher devices [15].

1) Core Bluetooth Framework for iOS: The Core Bluetooth framework provides the classes needed for your iOS and Mac apps to communicate with devices that are equipped with BLE wireless technology. For example, your app can discover, explore, and interact with low energy peripheral devices, such as heart rate monitors and digital thermostats. As of OS X v10.9 and iOS 6, Mac and iOS devices can also function as BLE peripherals, serving data to other devices, including other Mac and iOS devices [15].

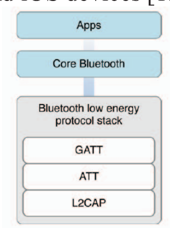


Figure 5: Bluetooth Stack [xi]

Key Terms	Description
Generic Attribute Profile (GATT)	The GATT profile is a general specification for sending and receiving short pieces of data known as "attributes" over a BLE link. All current Low Energy application profiles are based on GATT.
Attribute Protocol (ATT)	GATT is built on top of the Attribute Protocol (ATT). This is also referred to as GATT/ATT. ATT is optimized to run on BLE devices. To this end, it uses as few bytes as possible. Each attribute is uniquely identified by a Universally Unique Identifier (UUID), which is a standardized 128-bit format for a string ID used to uniquely identify information.
L2CAP	L2CAP is used within the Bluetooth protocol stack. It passes packets to either the Host Controller Interface (HCI) or on a hostless system, directly to the Link Manager/ACL link.

The Core Bluetooth framework lets iOS app communicate with Bluetooth low energy devices. For example, our app can discover, explore, and interact with low energy peripheral devices, such as external temperature Bluetooth devices or Texas Instrument Sensor Tag[xii] , and even other iOS devices.

2) Core Bluetooth Framework for Android: Android 4.3 (API Level 18) introduces built-in platform support for Bluetooth Low Energy in the central role and provides APIs that apps can use to discover devices, query for services, and read/write characteristics. In contrast to Classic Bluetooth, BLE is designed to provide significantly lower power consumption. This allows Android apps to communicate with BLE devices that have low power requirements, such as proximity sensors, heart rate monitors, fitness devices, and so on. [16].

E. Streaming Analytics & Edge processing

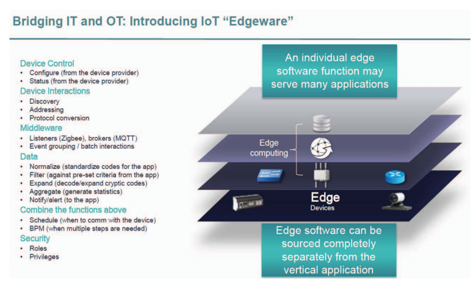
1) Streaming Analytic and Big Data:

Connected Dairy employs IoT World Forum proposed reference architecture. The reference architecture defines IoT stack on following seven levels:

1. Physical Devices & Controllers (The “Things” in IoT),
2. Connectivity (Communication & Processing Unit),
3. Edge Computing (Data Element Analysis & Transformation),
4. Data Accumulation (Storage)
5. Data Abstraction (Aggregation & Access)
6. Application (Reporting, Analytics, Control)
7. Collaboration & Processes (Involve People & Business Processes)

Edge Processing:

As the sensors emit streams of the data, the data processing can be accomplished near to the device/sensor or in the Cloud. For Connected Dairy, we have implemented processing closer to the Sensors – the so called “Fog or Edge Processing”. The primary use case to process at Edge level is real-time alert of “any urgent case” that Dairy needs to respond, or example, use cases such as a) precipitous drop in cattle temperature or b) sudden change in the ambient humidity.



Source: [xv]

The Edge processing triggers as soon as the mobile device receives data from sensors. The processing includes: sampling of the sensor event, time window evaluation, in-line rule processing, alert notification and delivery (if needed), persisting the event data in local memory, and post to on-premises or Cloud storage for reporting.

As part of edge processing, the architecture implements complex event processing (CEP). CEP enables connected dairy to handle:

- Real-time Pattern Recognition
- Real-Time Notification
- Closed Loop Notification

In order to preserve sensor events order and to handle any out of order event cases, the Connected Diary architecture handles the event “BLAST” gracefully. The BLAST stands for:

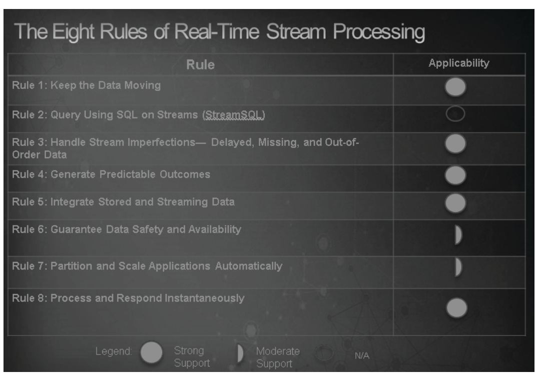
BLAST

- Burst Handling & Queuing
- Low Latency
- Availability
- Scalability
- Fault Tolerance

The robust nature of connected dairy real-time stream engine architecture is due to adherence to the principles of streaming processing [16]: [xvii]

- Rule 1: Keep the Data Moving
- Rule 2: Query Using SQL on Streams (StreamSQL
- Rule 3: Handle Stream Imperfections— Delayed, Missing, and Out-of-Order Data
- Rule 4: Generate Predictable Outcomes
- Rule 5: Integrate Stored and Streaming Data
- Rule 6: Guarantee Data Safety and Availability
- Rule 7: Partition and Scale Applications Automatically
- Rule 8: Process and Respond Instantaneously

Please note all the eight rules are not applicable for the connected dairy:

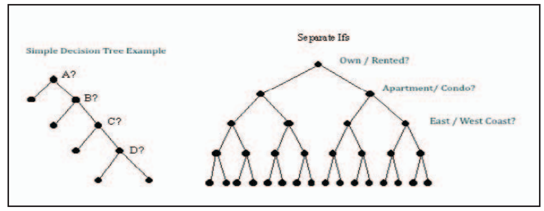

The Eight Rules of Real-Time Stream Processing

Rule	Applicability
Rule 1: Keep the Data Moving	●
Rule 2: Query Using SQL on Streams (StreamSQL)	○
Rule 3: Handle Stream Imperfections— Delayed, Missing, and Out-of-Order Data	●
Rule 4: Generate Predictable Outcomes	●
Rule 5: Integrate Stored and Streaming Data	●
Rule 6: Guarantee Data Safety and Availability	◐
Rule 7: Partition and Scale Applications Automatically	◐
Rule 8: Process and Respond Instantaneously	●

Legend: ● Strong Support ◐ Moderate Support ○ N/A

F. Machine Learning:

1) Decision Tree: A Decision tree algorithm is used to classify the attributes and decide the outcome of the class attribute. In order to construct decision tree both class attribute and item attributes are required. A Decision tree is a tree-like structure where the intermediate nodes represent attributes of the data, leaf nodes represent the outcome of the data, and the branches hold the attribute value. Decision trees are widely used in the classification process because *no domain knowledge is needed to construct* the decision tree [17]. The following figure shows simple decision trees.



The primary step in the decision tree algorithm is to identify the root node for the given set of data. Multiple methods exist to decide the root node of the decision tree. Information gain and Gini impurity are the primary methods used to identify the root node. Root node plays important role in deciding which side of decision tree the data falls into. Like every classification methods, decision trees are also constructed using the training data and tested with the test data.

Information Gain: Information gain is used to the root node and the branch nodes in the decision tree. Information gain is calculated using entropy and information. Entropy is calculated using the following formula [8].

$$Info(D) = -\sum_{i=1}^{m} p_i \log_2(p_i)$$

Information of the attribute is calculated using the following formula.

$$Info_A(D) = \sum_{j=1}^{v} \frac{|D_j|}{|D|} \times Info(D_j).$$

$$Gain(A) = Info(D) - Info_A(D).$$

Information gain of an attribute is the difference between entropy and the information of that attribute. The attribute with the highest information gain is the root node, and the next level nodes are identified using the next high information gain attributes. The algorithm and its pros and cons are listed below:

Algorithm

Step 1: Calculate the information gain for all the attributes
Step 2: Select the root node from the attribute list that has more information gain
Step 3: For each value of the root node
Step 4: Create a node for the attributes with next highest information gain.
Step 5: For each value of the nodes
Step 6: Create subset of training data for this node
Step 7: If all the values of class node are same, create a leaf node and stop
Step 8: Else go to step 5 and continue

Advantages	Disadvantages
• Simple and robust • Useful to predict the outcomes of future data • Little cleansing is enough to remove the missing values data • Useful for large data sets • Decision trees can handle both categorical and numerical data	• Possibility of creating complex decision trees for simple data • Replication problem makes the decision trees complex. So remove the replicated data before constructing a decision tree. • Pruning is required to avoid complex decision trees. • It is hard to find out the correct root node.

2) Memory Data Cubes: A data cube provides a multidimensional view of data and allows the pre-computation and fast accessing of summarized data. We have created Data Cube Memory Resident structures for accessing & processing streaming analytics.

3) Regression Analysis: In statistics, linear regression is an approach for modeling the relationship between a scalar dependent variable y and one or more explanatory variables denoted X. The case of one explanatory variable is called simple linear regression.

- **The simplest theoretical model: $Y = \beta_0 + \beta_1 X + \varepsilon$**
  - where β_0 and β_1 are parameters to be estimated
  - ε represents the influence of random factors (noise) on Y
- **Assumptions implicit in the model:**
  - the only "systematic" determinant of Y is X — there are no missing variables except "noise"
  - the relationship between Y and X is linear
  - ε has three very special properties--to be discussed soon
- **Define:**
  - b_0 and b_1 are estimates of β_0 and β_1 and
  - $Y^F = b_0 + b_1 X$, the forecast of Y for a given X
- **Note the difference**: $Y^F = b_0 + b_1 X$ and $Y = \beta_0 + \beta_1 X + \varepsilon$

•**Define**: the sum of squared residuals(errors),

$$SSR = \sum_{i=1}^{N}(Y_i - Y_i^F)^2 = \sum_{i=1}^{N} e_i^2$$

This is called SSE by SM and SSR by Excel

•Minimizing SSR (the total squared forecast error) yields the the formulas for calculating the so-called *least squares* estimates of β_0 and β_1---b_0 and b_1

•**Notice**: minimizing the total forecast error $\sum_{i=1}^{N} e_i$ will not yield desirable estimates of β_0 and β_1.

The Benchmark: Using $\bar{Y}$ to forecast Y

1) total squared forecast error, TSS

$$TSS = \sum_{i=1}^{N}(Y_i - \bar{Y})^2$$

2) mean squared forecast error, V(Y)

$V(Y) = TSS/(N-1)$ where N-1 is the number of degrees of freedom

3) mean forecast error, $\sqrt{V(Y)} = SD(Y)$

4) Sliding Window: Sliding-window analysis is a common way to analyze and identify trends in the stream data [18]. The term sliding-window analysis refers to a common pattern analysis of real-time and continuous data and uses rolling counts of incoming data to examine trending, temperature, and humidity.

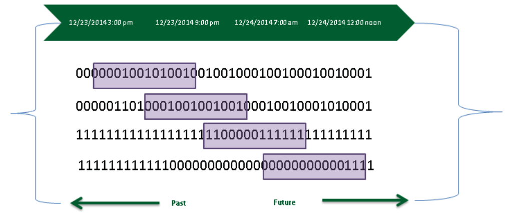


Figure 6: Time Window & Data Stream

III. SYSTEM OVERVIEW

Our goal is to provide a scalable cloud architecture which would provide a holistic approach for Connected dairy. The project will leverage the latest open source technologies to create a real-time big data processing platform. We employed Bluetooth Low Power Energy protocol to overcome some of the infrastructure limitations in Small to Medium size Dairies. We have taken extra considerations in saving battery by only interrogating the sensors at periods that Dairies would schedule.

A. System Architecture

The system architecture consists of five major parts or tiers: 1. Mobile Sensor Core, 2. DB Tier, 3. Service Tier, 4. UI Tier, and 5. Sensor Data Analytics Platform.

Mobile Sensor Core:

At the core of the system is Mobile interface that periodically connects and read the sensor values. The Mobile Device, Android or iOS, performs Edge processing before uploads the data back to the Cloud. The Edge processing involves checking the streams data from Dairy cattle, applying sliding window

Sensor Data Analytics Framework: Sensor Data Analytics Framework consists of a) Streaming Analytics, b) Machine Learning & Historical Analysis Component, c) Stream Ingestion Framework, and d) Data Store Component. The Dairy Sensor Data Analytics framework is based on two major Machine Learning Algorithms: a) Decision Tree and b) Sliding Window Streaming Analytics.

Database Tier: The Database Tier consists of NoSQL MongoDB. Given the Dairy sensor data signature differs from one BLE device to other (for instance, Sensor Tag vs. Third Party devices), our architecture aims to handle both structured and unstructured data. Second, for historical reporting purposes, architecture aims to save the post processed stream data. NoSQL Mongo DB is the ideal Database.

Service Tier: We have deployed our Service Tier on Tomcat Web Server with Representational State Transfer (REST)/ JavaScript Object Notation (JSON) Interface End Points. These Interface Points are accessed by Dairy Mobile User Devices and Dairy Management (Browser UI Client).

UI Tier: We have developed our Web UI with the help of HTML5, Bootstrap, and AJAX.

Mobile UI: The Mobile UI supports both Android (4.3 or higher) and iOS (7 or higher) frameworks. In Dairy settings, the Dairy User App connects to Sensor Tag using BLE, retrieves the data, applies edge rules, and stores the retrieved values in the Cloud by calling REST Interface functions.

Data Payload: We have designed JSON for Data communication between Mobile App and Service Tier.

{ "CAT_ID": "000023121ABM_20140101_TEL_HYD_DAIRY_00122", "TEMP_VAL": "39.23", "AMB_TEMP": "45.33", "HUMDITY": "60.9", "PRESSURE": "1015", "SESSION_TOKEN": "2333334", "GEO_LAT": "17.366", "GEO_LONG": "78.47", "MEASURE_TIME": "12/30/2014 -3:15 pm", "STATUS": "Ok", "ACTION":"None" }

Infrastructure: We have deployed the connected dairy on Elastic Cloud and intend to support multi-tent interface.

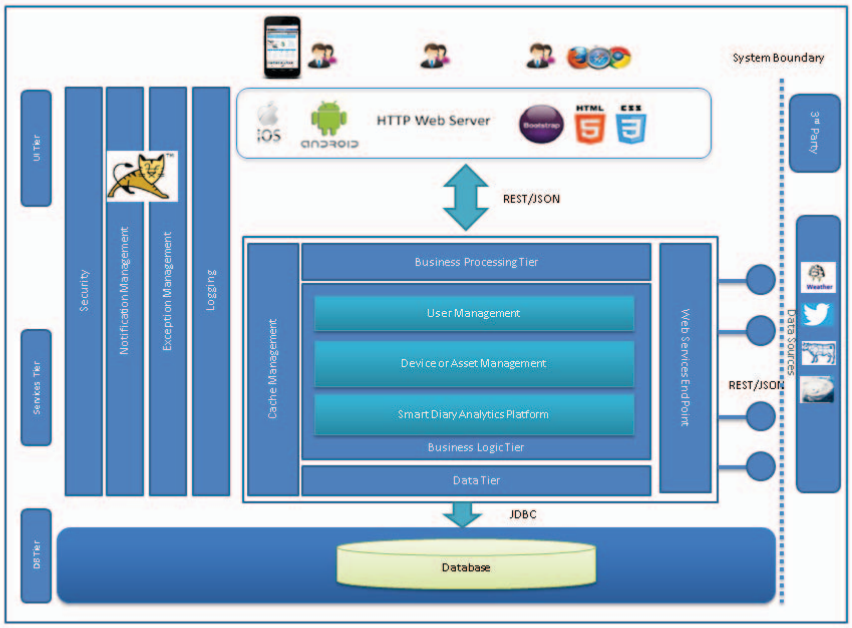


Fig. 1. **A Connected dairy IoT System Architecture**

B. System Function

1) Mobile Sensor Core: The primary goal of the Mobile Sensor Core is to retrieve Sensor tag values using BLE Protocol. One compelling reason is to increase working life of the Sensors, i. e., minimize battery drain.

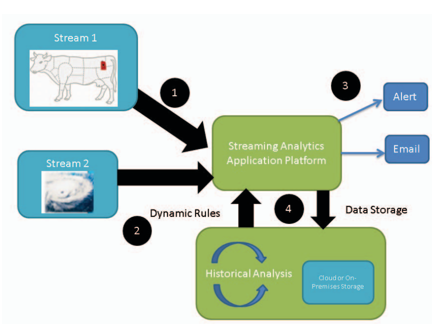


Figure 7: Sensor Call Sequence

Step 1: The Dairy User Mobile device connects to Sensor that is placed on the Cattle.
Step 2: Upon successful connection, the mobile device retrieves the value of core sensor values.
Step 3: In-line rule processing (sliding window) of retrieved data. Compares the values of most recent Sensor values and evaluates for any pre-configured rule violations.
Step 3b: Alerts the User upon any rule violation.
Step 4: Saves a copy on local phone and persists the Sensor values by uploading to the Cloud.

Here is the structure of JSON that persisted in database:

{ } JSON
CAT_ID : "000023121ABM_20140101_TEL_HYD_DAIRY_00122"
TEMP_VAL : "39.23"
AMB_TEMP : "45.33"
HUMDITY : "60.9"
PRESSURE : "1015"
SESSION_TOKEN : "2333334"
GEO_LAT : "17.366"
GEO_LONG : "78.47"
MEASURE_TIME : "12/30/2014 -3:15 pm"
STATUS : "Ok"
ACTION : "None"

Connecting Sensor Tag on iOS: Devices that implement the central role in BLE communication perform a number of common tasks, for example, discovering and connecting to available peripherals, and exploring and interacting with the data that peripherals have to offer. In contrast, devices that implement the peripheral role also perform a number of common, but different, tasks—for example, publishing and advertising services, and responding to read, write, and subscription requests from connected centrals.

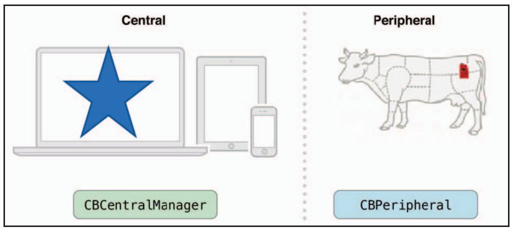


Figure 8: Connected dairy Interaction Design

The following steps are required to establish connection with Central & Peripheral:
a) Start up a central manager object (in our case, Dairy User App)
b) Discover and connect to peripheral devices (i.e., Sensor Tags) that are advertising
c) Explore the data on a peripheral device after you've connected to it
d) Send read and writes requests to a characteristic value (Sensor Tag values) of a peripheral's service
e) Subscribe to a characteristic's value to be notified when it is updated

Connecting Sensor Tag on Android: Android 4.3 (API Level 18) introduces built-in platform support for Bluetooth Low Energy in the central role and provides APIs that apps can use to discover devices, query for services, and read/write characteristics. In contrast to Classic Bluetooth, BLE is designed to provide significantly lower power consumption. This allows Android apps to communicate with BLE devices that have low power requirements, such as proximity sensors, heart rate monitors, fitness devices, and so on.

BLE Permissions:
To enable the application to support the Bluetooth features, we must add the Bluetooth Permission BLUETOOTH to the android manifest file. This gives the app the permission to interact with other Bluetooth devices such as requesting a connection, accepting a connection, and transferring data. The application also includes the BLUETOOTH_ADMIN permission so that it can initiate device discovery or manipulate Bluetooth settings

Figure 9: Connected dairy Android App Screen Capture

C. *Key Technologies*

1) *Memory Resident Data Cubes:*

In order to process and apply series of time window & time-sensitive rule algorithms, we have implemented persistent memory resident data cubes. First, as soon as the sensor values are retrieved, the mobile device saves the values on device internal storage – archive file. This way the architecture takes care of any intermittent network issues. Second, the collected sensor values are translated and inserted in to Data cube cells for analysis purposes.

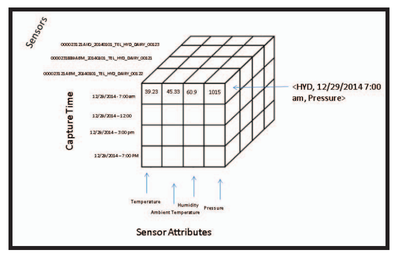


Figure 10: Data Cube Representation

The advantage of memory resident data cube is we can apply time series and Slide window analysis on it.

2) Streaming Analytics: We have designed Streaming Analytics on Android and iOS Mobile App level. Upon receiving Sensor Stream data, the Stream analytics component triggers. The Streaming process first load the new sensor values in a memory resident data store and applies slide window process to check any rule violation.
Here are the Stream evaluation steps:
1. Collect Data rea-time
2. Process the Streaming Data
3. Explore, Analyze, and Visualize
4. Action

During the analysis of the events for patterns detection, the architecture compares the sequence of sensors and loads the sensor event data in the data cube delimited by Sensor ID, Date & Time.

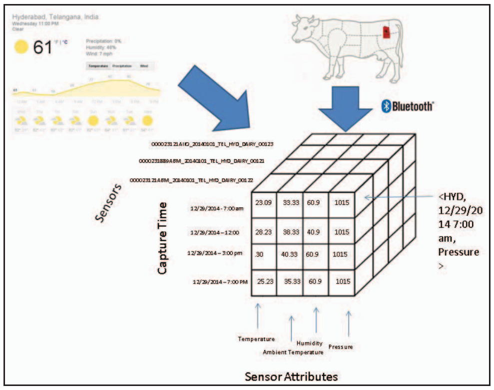


Figure 11: Sensor and Weather - Data Cube Representation

Upon loading the values into Data Cube, the sliding Window algorithm is applied to detect any anomalies or rule violation.

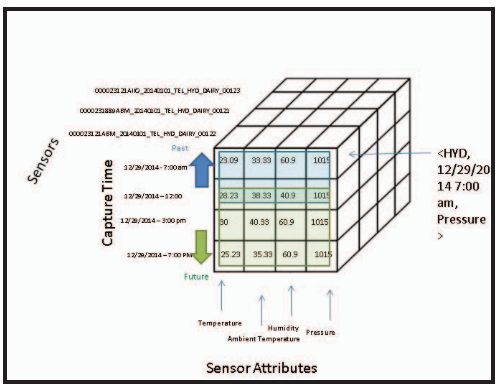


Figure 12: Data Cube Sliding Window

Once, streaming analysis is complete the sensor values are uploaded to Cloud on REST Interface.

3) *Pattern Detection and Predictive Analytics:*

Our architecture stores the Streaming analytics values back into the central database for reporting and trend analysis purposes. We have developed Predictive algorithms to help Dairy to pro-actively handle any operational policies or infrastructure needs. In essence, Predictive Analytics module help deliver operational intelligence that help reduce cost of maintenance.

In order to develop Predictive Analytics, we have taken nine months Dairy training dataset and developed model on top of it. Given our analysis, we are successfully predicting the Dairy temperatures with the accuracy of 90% and above

Our model has following attributes:

<u>Sensor ID:</u> Dairy cattle Sensor values. During the system configuration, each Dairy cattle assigned an unique Sensor.
<u>Date:</u> Date of the Sensor Collection.
<u>Time:</u> The time Dairy personnel retrieved the values. Typically, in our observation, the Dairy personnel retrieve the values before dawn, during noon, evening, and night. Our architecture does not assume any time and frequency of retrieval.
<u>Local Temperature:</u> Prevailing temperature of the location. Typically, the local temperature is retrieved by call 3rd party Weather REST Services.
<u>Body Temperature:</u> Cattle Body Temperature. TI Sensor Tag provides the Object temperature. In our Connected Dairy, we retrieve the Body Temperature of the Cattle by placing on sensor on the cattle. Our architecture does not restrict to the Object temperature. We can interrogate other modes of temperature retrieval too. For instance: rectal temperature.
<u>Ambient Temperature:</u> The Dairy settings temperature.

<u>Pressure:</u> The barometer pressure
<u>Humidity:</u> The humidity of local Weather. Typically, we retrieve Humidity by calling 3rd party REST Calls
<u>WaterInTake</u>: Observed Water intake of cattle
<u>FoodInTake</u>: Observed Food supplied to Cattle
<u>FeedType</u>: Type of Feed provided
<u>AC On</u>: Air conditioning setting (ON/OFF)
<u>FanOn</u>: Fan is on (On/OFF)
<u>MilkOutput</u>: Milk produced by the cattle
<u>Medication</u>: any medication provided to cattle
<u>Medical conditions</u>: Cattle health status, Normal, Sick with pneumonia ect.
<u>Vitamins</u>: Any Vitamins provided

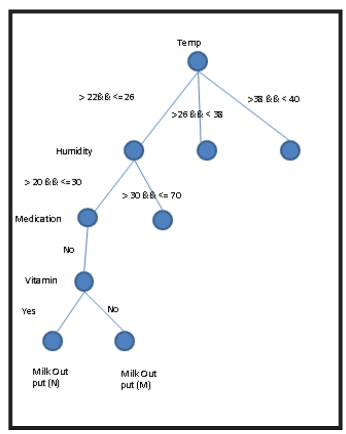


Figure 13: Training Dataset (Simple DT) Not holistic view

Here is attached sample data:

Sensor ID	Date	Time	BodyTemp	AmbientTemp	Humadity	WeatherTemp	Pressure	FoodInTake	WaterInTake	AC	Fan	WaterTreatment	MilkOutput
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/28/2014	7	23	33	60	20	1015	30	30	0	1	0	10
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/28/2014	12	27	37	50	25	1015	40	25	1	1	0	10
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/28/2014	15	30	40	50	29	1015	40	5	1	1	1	10
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/28/2014	20	25	30	70	18	1015	50	10	0	1	0	10
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/29/2014	7	20	33	60	18	1015	30	30	0	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/29/2014	12	24	37	50	23	1015	50	25	1	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/29/2014	15	27	40	50	22	1015	40	5	1	1	1	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/29/2014	20	22	30	70	20	1015	50	10	0	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/30/2014	7	23	33	60	20	1015	30	30	0	1	0	10
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/30/2014	12	27	37	50	25	1015	40	25	1	1	0	10
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/30/2014	15	30	40	50	29	1015	40	5	1	1	1	10
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/30/2014	20	25	30	70	18	1015	50	10	0	1	0	10
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/31/2014	7	20	33	60	18	1015	30	30	0	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/231/2014	12	24	37	50	23	1015	50	25	1	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/31/2014	15	27	40	50	22	1015	40	5	1	1	1	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	12/31/2014	20	22	30	70	20	1015	50	10	0	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	1/1/2015	7	20	33	60	18	1015	30	30	0	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	1/1/2015	12	24	37	50	23	1015	50	25	1	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	1/1/2015	15	27	40	50	22	1015	40	5	1	1	1	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	1/1/2015	20	22	30	70	20	1015	50	10	0	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	1/2/2015	7	20	33	60	18	1015	30	30	0	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	1/2/2015	12	24	37	50	23	1015	50	25	1	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	1/2/2015	15	27	40	50	22	1015	40	5	1	1	1	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	1/2/2015	20	22	30	70	20	1015	50	10	0	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	1/3/2015	7	20	33	60	18	1015	30	30	0	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	1/3/2015	12	24	37	50	23	1015	50	25	1	1	0	11
000023121ABM_20140101_TEL_HYD_DAIRY_00122	1/3/2015	15	27	40	50	22	1015	40	5	1	1	1	11

Figure 14: Sensor Captured Values - Location Vishakpatanam India

We have run the regression analysis on the supplied data and developed model equation. Model Regression Analysis and Model Equation:

MilkOutput = 41 + (-0.33) * Body Temp + (-0.272) * Ambient Temp + (-0.196) * Humidity + (0.44) * Food Intake + (0.55) * Vitamins + (-0.3) * Medications + (0.01) * AC On + (0.33) * Fan On

Figure 15: Regression Model Equation (Milk output vs. Dairy parameters)

Please note: the above model equation is developed based on nine months of datasets from a Dairy in Vishakhapatnam, Andhra Pradesh, India. The equation is not a generic one and will not be applicable to other Dairy settings. Nonetheless, the input parameters—temperature and humidity—have considerable impact on milk productivity [3].

SUMMARY OUTPUT

Regression Statistics	
Multiple R	1
R Square	1
Adjusted R Square	0.75
Standard Error	3.68379E-16
Observations	27

ANOVA

	df	SS	MS	F	gnificance F
Regression	11	5.62963	0.511785	6.91E+30	3.4E-228
Residual	20	2.71E-30	1.36E-31		
Total	31	5.62963			

	Coefficients	andard Err	t Stat	P-value	Lower 95%	Upper 95%	ower 95.0%	pper 95.0%
Intercept	23.71929825	2.66E-15	8.91E+15	1.8E-307	23.7193	23.7193	23.7193	23.7193
Time	0.052631579	1.25E-16	4.2E+14	6.3E-281	0.052632	0.052632	0.052632	0.052632
BodyTemp	-0.333333333	7.2E-17	-4.6E+15	8.8E-302	-0.33333	-0.33333	-0.33333	-0.33333
AmbientTemp	0	0	65535	#NUM!	0	0	0	0
Humadity	-0.085964912	3.17E-17	-2.7E+15	#NUM!	-0.08596	-0.08596	-0.08596	-0.08596
WeatherTemp	-2.29445E-17	4.91E-17	-0.46721	0.645403	-1.3E-16	7.95E-17	-1.3E-16	7.95E-17
Pressure	0	0	65535	#NUM!	0	0	0	0
FoodInTake	2.07489E-17	3.58E-17	0.580344	#NUM!	-5.4E-17	9.53E-17	-5.4E-17	9.53E-17
WaterInTake	-0.042105263	3.08E-17	-1.4E+15	3.5E-291	-0.04211	-0.04211	-0.04211	-0.04211
AC	0	0	65535	#NUM!	0	0	0	0
Fan	0	0	65535	#NUM!	0	0	0	0
WaterTreatment	0	0	65535	#NUM!	0	0	0	0

Figure 16: Model Regression Equation

To quantify the relationship between heat stress and dairy production the authors used a commonly accepted and tested model by Berry et al. (1964). This model uses both temperature and humidity to calculate the loss in normal production levels of a given dairy cow. The equation is displayed below [3]:

$$MPD = -1.075 - 1.736 \cdot NL + 0.02474 \cdot NL \cdot THI$$

Where
NL= normal production levels, kg/(cow x day)
THI = temperature humidity index
MPD = the absolute decline in milk production, kg/(cow x day)

4) ***Machine learning with high performance:*** An in-memory cluster computing platform is used that increases performance versus traditional Hadoop deployment. This allows our platform to load data into a cluster's memory and query it repeatedly making it suitable for different machine learning algorithms. This allows to process data faster, and thus helps in scaling the application.

IV. System Design And Implementation

A. *System Component Architecture Design*

The System contains Mobile Streaming Analytics, Analytics at Cloud, and Machine Learning Algorithm for Forecasting and Recommendation.

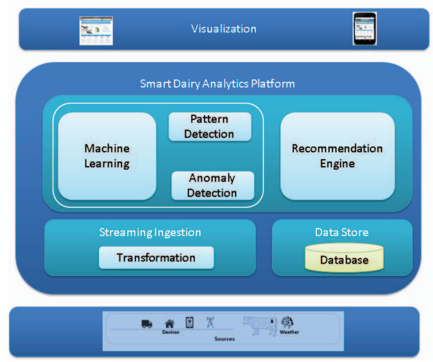


Figure 17: Analytics Platform

Connected dairy Analytics Platform: Analytics platform consists of two core engines: a) Machine Learning and b) Recommendation Engine.

<u>The Machine Learning and Rules processing</u> are performed both at the Edge level and in the Cloud. The Edge Rules will alert the Dairy personnel and the rules engine in the Cloud will help provide historical analysis.

```
If Body Temp is between 22 degree C and 26 degree C then
        if Rel_Humidity is between 20 per to 30 per then
                // No Heat Stress
                // No Need to perform extra steps
        else if Rel_Humidity > 30 and Rel_Humidity < 70 Per then
                // Moderate Heat Stress
                // Starting decresing Thremostat to get into lesser Humidity
                // no harm feeding more water....
        else
                // Do nothing
        end if

        -
        -
        -
        -
        -
        -
        -
Else If Body Temp is between 38 degree C and 40 degree C then

        if Rel_Humidity is between 20 per to 80 per then
                // Severe Heat
                // Call Vet.
        else if Rel_Humidity is between 90 per to 100 per then
                // Very deadly condition
                // Please check cattle immediately
                // Send Alert
                // Call Emergency
        else
                // Do Nothing
else
        // Do Nothing

End if
```

Figure 18: Rule Editor

Pattern Detection: Using Machine Learning and Pattern detection, we can correlate the relation between several Dairy optimization factors. For instance, the impact of medication and vitamin intake is on the Dairy cattle Milk output. The impact of Seasonal influences on the Health & Milk output.

Historical Analysis: Historical analysis includes the influence of season factors such as HS and pneumonia on Dairy cattle output. Moreover, our architecture ties the weather forecast, i.e. weather patterns in coming weeks, to the historical observed behavior on Dairy cattle. For instance, if a Dairy cattle shows illness related symptoms due to sudden temperature changes, our architecture provides actionable insights to Dairy personnel by looking into any sudden Weather changes in near future and immediately alerting Dairy personnel if there any such events.

Recommendation System: Our recommendation system helps Dairy management to handle: Food optimization, Medication Suggestions, Optimal Dairy settings per cattle, and Disease Management.

We have developed our HS rule recommendations based on the findings of Davison *et al*. [1996] who found that a cow's milk yield starts decreasing when the Temperature Humidity Index reaches 72. They investigated three separate management techniques and found their resulting impact on milk yields. The results are shown in the following figure:

Management	Poor	Average	Good	Best
Cooling Strategy	Nil	Some Shade	Shade at Feed	Shade & Sprinklers
THI Threshold	72	74	76	78

Figure 19 THI thresholds leading to stress under different management techniques (Davison et al., 1996)[xviii].

<u>**Other Components:**</u>

1) ***Security:*** Our system shall support RBAC – Role Based Access Support.

2) ***Web services and APIs:*** For providing web services, we will build up web services and APIs for getting results from services layers. This component will be on the layer of communication.

3) ***User Interface:*** Our main Users are Dairy operation and management staff. The Operation Staff will be equipped with Mobile Devices that can read the values from the Sensors. The Management individual will have Dashboard view that provides more insights into day-to-day operations and, importantly, provides forecasting based on weather patterns.

In order to forecast operational insights, we would retrieve weather forecast from third party Connectors - Weather engines.

B. System Interface and Connectivity Design

In this section we will discuss the system interface and connectivity design of the analytics service platform. The end-user or customer logs into the system with a device through the internet and then browses the web or clicks an app. Then the service request will be sent to the Web server and Web server communicate with the recommendation engine. The recommendation engine uses the data and algorithm stored in the database to perform real-time recommendation to the end users.

a) ***C. Recommendation Engine Process Flow and Innovations:*** The whole process flow can be divided into step key steps: Pre-process, Machine learning process and post-process. In each of these three steps, there are several innovations.

V. A CASE STUDY

We have tested and deployed the connected dairy in one dairy in Visakhapatnam, Andhra Pradesh, India. Our initial analysis reveals that the customer, Dairy management, and operational personnel hugely benefited from the findings of the Connected Dairy. We have received extensive requests to provide historical analysis, recommendations, and operation reports. Attached screen captures of connected diary.

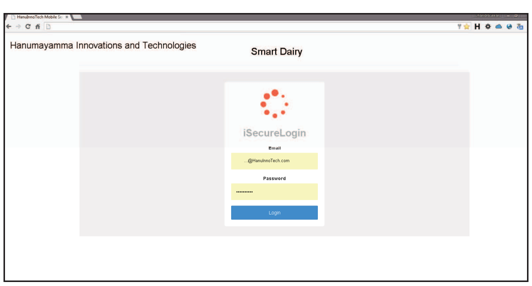


Figure 20: Login

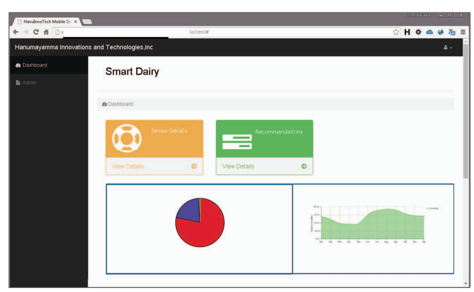


Figure 21: Connected dairy Dashboard

Figure 22: Mobile UI

VI. CONCLUSION AND FUTURE WORK

This paper presented a novel approach to addressing the critical needs of Dairy industry. We strongly believe that the Connected Dairy will yield huge operational efficiencies, cost savings, and actionable insights to address Dairy cattle related critical issues. Connected dairy, importantly, is a data enabled insightful tool that facilitates better management of Dairy activities. Finally, connected dairy provides forecasting insights that provides window of time opportunity to dairy operational management so that they can better plan to handle any un-expected weather related abnormalities, Dairy cattle health and emergencies.

References

[1] Ian Ohnstad, "Managing Heat Stress in Dairy Cows", [Online], Available: http://www.nadis.org.uk/bulletins/managing-heat-stress-in-dairy-cows.aspx

[2] Dave Warner , "Keeping cows cool critical to dairy industry as climate warms", [Online], Available: http://www.pennlive.com/midstate/index.ssf/2013/04/keeping_cows_cool_critical_to.html

[3] Juliano Calil,Alexander Silvester,Katharina Stelzl, Christopher Wissel-Tyson, "The Effect of Climate Change on the Production Costs of the Dairy Industry in the United States", [Online], Available: http://www.bren.ucsb.edu/research/2012Group_Projects/documents/gotmilk_report.pdf

[4] Jeffrey F. Keown and Richard J. Grant, "How to Reduce Heat Stress in Dairy Cattle", October 1993, [Online], Available: http://extension.missouri.edu/p/G3620

[5] Legates JE1, Farthing BR, Casady RB, Barrada MS., "Body temperature and respiratory rate of lactating dairy cattle under field and chamber conditions. ", 1991 Aug;74(8):2491-500, [Online], Available: http://www.ncbi.nlm.nih.gov/pubmed/1918528

[6] Jon Marcus, "How Cows Are Becoming Smart Connected Products", June 2014, [Online], Available: http://www.forbes.com/sites/ptc/2014/06/18/how-cows-are-becoming-smart-connected-products/

[7] Abi Grogan, " Internal smart pill to save the dairy industry $10bn", June 18, 2014, [Online], Available: http://eandt.theiet.org/news/2014/jun/vital-herd-smart-pill.cfm

[8] Gartner, "Gartner Says the Internet of Things Installed Base Will Grow to 26 Billion Units By 2020", December 12, 2013, [Online], Available: http://www.gartner.com/newsroom/id/2636073

[9] IBM, "Internet of Things", http://www-01.ibm.com/software/info/internet-of-things/

[10] Michael E. Porter and James E. Heppelmann, "How Smart, Connected Products Are Transforming Competition", Harvard Business Review, November 2014 Issue, Available: https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition

[11] Beecham Research, "M2M/IoT Sector Map", [Online], Available: http://www.beechamresearch.com/article.aspx?id=4, Accessed: 01//05/2015

[12] Apple Developer Page, "About Core Bluetooth", [Online], Accesed: 12/27/2014, Available: https://developer.apple.com/library/ios/documentation/NetworkingInternetWeb/Conceptual/CoreBluetooth_concepts/AboutCoreBluetooth/Introduction.html#//apple_ref/doc/uid/TP40013257-CH1-SW1

[13] TI Sensor Tag, Accessed 12/25/2014,[Online], Available: http://processors.wiki.ti.com/index.php/Simplelink_SensorTag

[14] Wikipedia, "Bluetooth low energy", [Online], Accessed: 12/28/2014, Available: http://en.wikipedia.org/wiki/Bluetooth_low_energy

[15] Apple Developer Page, "About Core Bluetooth", [Online], Accesed: 12/27/2014, Available: https://developer.apple.com/library/ios/documentation/NetworkingInternetWeb/Conceptual/CoreBluetooth_concepts/AboutCoreBluetooth/Introduction.html#//apple_ref/doc/uid/TP40013257-CH1-SW1

[16] Google Developer Page, "Bluetooth Low Energy", [Online], Accesed: 12/27/2014, Available: https://developer.android.com/guide/topics/connectivity/bluetooth-le.html

[17] Jiawei Han ,Micheline Kamber and Jian Pei, *Data Mining: Concepts and Techniques*, Morgan Kaufmann; 3 edition (July 6, 2011)

[18] Anand Rajaraman and Jeffrey David Ullman , *Mining of Massive Datasets*, Cambridge University Press (December 30, 2011)

---

[i] http://www.pennlive.com/midstate/index.ssf/2013/04/keeping_cows_cool_critical_to.html
[ii] http://www.forbes.com/sites/ptc/2014/06/18/how-cows-are-becoming-smart-connected-products/
[iii] http://extension.missouri.edu/p/G3620
[iv] http://www.nadis.org.uk/bulletins/managing-heat-stress-in-dairy-cows.aspx
[vv] https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition
[vii] BLE http://www.bluetooth.com/Pages/Bluetooth-Smart.aspx
[ix] http://www.ti.com/ww/en/wireless_connectivity/sensortag/index.shtml?INTC=SensorTag&HQS=sensortag-bt
[x] http://processors.wiki.ti.com/index.php/Simplelink_SensorTag
[xi] https://developer.apple.com/library/ios/documentation/NetworkingInternetWeb/Conceptual/CoreBluetooth_concepts/AboutCoreBluetooth/Introduction.html#//apple_ref/doc/uid/TP40013257-CH1-SW1
[xii] Sensor Tab - http://www.ti.com/ww/en/wireless_connectivity/sensortag/index.shtml?INTC=SensorTag&HQS=sensortag-bt
[xv] https://s3.amazonaws.com/cdn.iotwf.com/breakouts/2014/H-ARC-01_Cisco-Intel-IBM_FINAL.pdf
[xvii] http://complexevents.com/wp-content/uploads/2006/07/StreamBaseEightRulesWhitepaper.pdf
[xviii] http://www.bren.ucsb.edu/research/2012Group_Projects/documents/gotmilk_report.pdf