Transportation and distribution (T&D) of fresh food products is a substantial and increasing part of the economic activities throughout the world. Unfortunately, fresh food T&D not only suffers from significant spoilage and waste, but also from dismal efficiency due to tight transit timing constraints between the availability of harvested food until its delivery to the retailer. Fresh food is also easily contaminated, and together with deteriorated fresh food is responsible for much of food-borne illnesses.

The logistics operations are undergoing rapid transformation on multiple fronts, including infusion of information technology in the logistics operations, automation in the physical product handling, standardization of labeling, addressing and packaging, and shared logistics operations under 3rd party logistics (3PL) and related models. In this article, we discuss how these developments can be exploited to turn fresh food logistics into an intelligent cyberphysical system driven by online monitoring and associated operational control to enhance food freshness and safety, reduce food waste, and increase T&D efficiency. Some of the issues discussed in this context are fresh food quality deterioration processes, food quality/contamination sensing technologies, communication technologies for transmitting sensed data through the challenging fresh food media, intelligent management of the T&D pipeline, and various other operational issues. The purpose of this article is to stimulate further research in this important emerging area that lies at the intersection of computing and logistics.

CCS Concepts: • Hardware → Sensor applications and deployments; Sensor devices and platforms; • Networks → Sensor networks;

Additional Key Words and Phrases: Food supply chain, communication infrastructure, physical Internet, food quality sensing, data analytics, security and privacy

ACM Reference format:

1 INTRODUCTION
Transportation and distribution (T&D) of fresh food is a huge and growing enterprise due to worldwide sourcing of products and increasing recognition of its benefits. Yet, fresh food transportation and distribution (T&D) is very inefficient in terms of substantial spoilage and wastage, dismal
Fig. 1. Food losses categorized by the types and stages of the food pipeline. The four columns (of a commodity type) denote the estimated waste percentages in Europe, North America, Asia, and sub-Saharan Africa, respectively. Notice that the bars show the food losses in percentage of what enters at each step. The data for generating this figure are obtained from Reference [77].

utilization of T&D capacity, and large energy and carbon footprint for fuel and cooling. The inefficiency results primarily from the need to have the transportation capability (e.g., empty containers, trucks, drivers, loaders) ready when the fresh food becomes available (e.g., harvested) and there are contracts to be fulfilled. This often results in proactive movement/positioning of empty assets often known in the industry as “shipping air” and can substantially reduce the usage of available transport capacity and thus lead to poor efficiencies. The increasing customer expectations of food in prime condition and without any treatment exacerbates the problem. Current T&D efficiency estimates are in 10%–20% range, and food spoilage/waste estimates in ~12% range, although the total food wastage including at the source and by consumers exceeds 40%, or $165B per year in the US alone. Figure 1 shows the percentage of food loss at different stages (production, post-harvesting, processing, distribution, and consumption) of the food pipeline, at different regions. A significant portion of the fresh food losses occur from the post-harvesting to consumption stage due to poor and uninformed handling/sorting during transportation and storage, lack of adequate ambient control, and so on (as summarized in Figure 2), which can be largely avoided [16, 22]. The actual estimates of food losses in the supply chain vary quite a bit. In Reference [129], the losses in the US supply chain are listed as 2%–30%, but closer to 15%. In Figure 2, the items in black boxes can also be considered as related to sensing/communications infrastructure in that such an infrastructure can detect the problem early and potentially avoid waste by altering delivery schedule or destination. For example, in India a large amount of food waste happens immediately after harvest due to the lack of proper storage facilities. In Reference [137], the authors have described the use of zero-energy cool chambers, which significantly increases the shelf life of the stored fruits and...
vegetables as depicted in Figure 3. Figure 4 shows the effect of ambient control on the chemical changes of tomatoes and cabbages in three retail outlets after one week of storage, which also shows a significant amount of quality gain due to ambient control.

In addition, fresh food is almost entirely responsible for large annual incidence of food-borne illnesses in the US and elsewhere. The CDC maintains extensive data on this subject [48], stating that "each year roughly 1 in 6 Americans (or 48 million people) get sick, 128,000 are hospitalized, and 3,000 die of foodborne disease." Of these, about 46% of cases are due to produce, 20% dairy/eggs, and the rest to meat/fish. Such food poisoning or contamination happens in different stages of the food production chain—from farms to the cutting boards in the kitchen. Tracing back the origin of such food-borne disease is quite challenging due to lack of comprehensive recording of the handling of the food in the entire supply chain [128]. Some of the worst food-borne illness outbreaks in the USA are reported in Reference [24], which shows that such contamination can happen in any stage of the food pipeline. Although the outbreaks contribute to less than 5% of
Fig. 4. Percentage change of Ascorbic acid in tomatoes and cabbages from three outlets at ambient and optimum conditions after seven days [115]. In case of tomatoes, the ambient and optimum conditions are 22–25°C, 52%–57% relative humidity and 10–12°C; 92%–95% relative humidity, respectively. In case of cabbages, these numbers are 22–25°C, 52%–55% relative humidity and 0°C, 91%–97% relative humidity, respectively.

the lab-confirmed foodborne illnesses [48], this would amount to 2.4M people (on 48M base) that could potentially be prevented from getting sick in the USA with better contamination detection.

Fortunately, there are many technologies coming together to enable substantial automation in operations, real-time quality/contamination monitoring, and agile logistics to adapt operations to reduce inefficiencies and food waste. Thus, the fresh food T&D can be considered as an emerging cyber–physical–human system with substantial advances and emerging challenges in all three components (cyber, physical, and human). The purpose of this article is to discuss these technologies, particularly relative to the cyber aspect. This includes food quality sensing, robust low-power communications to report sensed data in a very challenging environment, real-time monitoring and data collection, and online analytics to drive agile logistics operations. In particular, we envision an entire system for online monitoring of quality/contamination at all stages of distribution, communication of this data to a (logically) central location for analytics that would enable proactive mechanisms that enhance food freshness and safety, reduce food waste, increase T&D utilization and energy efficiency, while taking care to manage the sensing infrastructure to maximally reuse the sensors, extend their battery lifetime, eliminate malfunctioning sensors, and so on.

The article will comprehensively discuss mechanisms for monitoring perishability/contamination in the fresh food T&D pipeline. Efficient sensing and communications are central to this vision and pose several challenges because of the need for near-field communications in an extremely challenging environment of tissue/saline media. We advocate different communication techniques that appear to be particularly suited for this environment and discuss their pros and cons. We propose to collect network-wide data and exploit it to help reduce perishable food waste as well as proactively identify contamination origin in case of a food-borne illness to ensure quick response. Intelligent data analytics can also be used for better cooling management as well as building smarter distribution to reduce waste.

The outline of the article is as follows: Section 2 discusses the overall context of the research and its larger vision. Section 3 describes several factors that influence the food quality during different stages of food transportation. Section 4 describes different perishability metrics and shelf life modeling–related researches for fresh food. Different food sensing technologies available in the market are discussed in Section 5. Section 6 describes several communication challenges that arise while building a multi-hop sensor network in food (aqueous or tissue) medium. Section 7 describes the overall sensing and communication infrastructure of our proposed framework.
Section 8 describes the different data analytics and control that can potentially revolutionize the food logistics efficiency. Section 9 describes how the proposed infrastructure can assist the supply chain in reducing food waste or quickly locating the origin of contamination. The article is concluded in Section 10.

2 MOTIVATION AND BACKGROUND

The overall architecture of a fresh food T&D pipeline is shown in Figure 5. Foods from farmlands are taken to the packing centers where they are sorted, packed, and palletized for delivery to a nearby distribution center. The pallets may pass through multiple distribution centers on their way to the last-mile distribution center where the pallet may be unbundled for delivery to the retailers. We assume that some sensing and communication modules are placed in some of the boxes/pallets that are in transit or stored in some warehouse/distribution/retailing centers. These modules sense the quality/contamination parameters and deposit them in a local hub by forming a local communication network. These local hubs send this information to an Analytics & Operations Center (AOC) via long-range cellular communication. The important aspects of a smart T&D supply chain are: (a) to devise mechanisms for online sensing/communication of the food packages throughout the supply chain, (b) centralized collection of perishability and contamination data in the AOC, and (c) analysis of this data to enable intelligent and proactive distribution that enhances food freshness and safety, reduces food waste, and increases T&D utilization and efficiency.

2.1 Why Is Fresh Food Logistics Hard?

The transportation and distribution (T&D) supply chain requires an intricate interplay of many processes to efficiently carry goods from the producers (e.g., factories and farmers) to the supply chain customers such as retailers. The supply chain invariably includes a hierarchy of distribution centers to enable not only the store-and-forward transit but also bundling/unbundling of packages, quality monitoring, inventory control, and other functions. The coordination of storage with the incoming and outgoing transit is a crucial function both for space and quality management.
Another crucial aspect is the proactive positioning of resources such as empty containers, empty trucks, available drivers, and so on, at or near the sources of the goods so they can be available when needed. Often the deliveries are governed by contracts that have financial implications if violated. The problem becomes even more complex with fresh food both because of the perishability-related timing constraints and uncertainties in the harvesting/availability of the food for transport. With increasing variety of fresh food, it also becomes difficult to fill a truck with one product type. Sharing a truck among different products, each with its own perishability characteristics, coupled with multiple pick-up and drop-off points make the problem extremely complex. Note that carrying multiple product types may require different types of containers, each designed with certain cooling, ventilation, humidity, and vibration characteristics. These issues are often handled by overprovisioning of resources and avoiding product mixing or multipoint pick-up or drop-off. The result is a sweet spot that has both low efficiency (often less than 15%) and still a significant loss of quality, including outright spoilage.

Yet another factor in the low efficiency (and therefore high cost) of T&D supply chain is the prevalence of private logistics. Large, integrated companies such as Walmart, Target, and so on, tend to create their own private logistics network in the hopes of better control over operations, better privacy, and other perceived advantages. However, private logistics means poorer utilization of resources such as warehouse/truck space, more difficult reverse logistics (i.e., positioning of trucks and containers), and so on. The concept of 3rd party logistics (3PL) and its derivatives have tried to ameliorate this situation by outsourcing logistics to a third party that works with multiple clients and provides them with the desired cost vs. service quality tradeoffs. 3PL has grown consistently and now accounts for more than 54% of the transportation. Also, its variants (e.g., US Postal Service and UPS sharing delivery runs) also continue to materialize. While 3PL can, in theory, achieve higher efficiencies and lower costs by virtue of sharing of resources and volume advantages, implementing a large variety of service levels for different clients with different types of products and needs becomes even more challenging. It is also worth noting that the traditional practice of inflexible long-term contracts that the client typically likes to have only makes things more difficult in a 3PL environment with multiple parties involved.

2.2 Recent Technological Developments

It should be no surprise that logistics operations have become increasingly automated. Since dealing with small retail-size boxes is difficult and unnecessary in the T&D pipeline, such boxes are invariably palletized in a rectangular space—usually a few meters across in length, width, and height. The palletization/depalletization is relatively easy to automate and is already widely practiced, as shown in Figure 6. Other common automation tasks include loading/unloading of pallets into carriers (trucks, railcars, etc.), sorting or rearrangement of items, driverless transport in local areas, and so on. These capabilities that apply to a large variety of industries are often referred to as the "Industry 4.0” initiative, and the automation will continue to increase at a rapid pace. The automation is a crucial piece in the intelligent management of the T&D pipeline; for example, with automated palletization, we can assume that the boxes in the pallet are always arranged in a consistent pattern and the palletizer can record the relative order of the boxes. This, in turn, simplifies identification and handling of boxes within a pallet.

The automation crucially depends on the electronic tagging/identification of the items that it handles. In the logistics space, this means we need electronic tags at all levels, including individual retail boxes, pallets of boxes, containers, forklifts, sorters, carriers (trucks, railcars, etc.), warehouse rooms/sections, and so on. For example, if the palletizer can read the tags of all the boxes that go into a pallet and keeps track of the order in which the boxes go into the pallet, we have complete information about which box is where in a pallet. At a minimum, the tag should provide a
unique address to identify each item, whose other attributes could then be looked up in appropriate databases. Self tracking of items requires dynamic labelling; for example, as the pallet moves to the next node (a warehouse or distribution center), the tag reading that is done on arrival could also update a part of the tag so the pallet always carries its own location. The updated part could be a small sequence number or the label of the node itself. It is also possible to simply append the sequence number of the next node and thereby retain the entire history of the pallet movement. (One could also append the label of the next node directly; however, this may require too many bits in the tag.)

The most common electronic tagging is done using RFID (radio frequency identification) [46, 152]. RFID tags are small transponders that are characterized by unique IDs that are exchanged while wirelessly responding to queries to the RFID readers. RFID tags are of three types: passive, semi-passive, and active. The passive tags consist of a micro-chip and an antenna. These tags do not have an internal power source or battery. The reader generates an electro-magnetic field that induces current into the tag’s antenna and powers up the chip. Once activated, the tags respond to the readers by using backscattering to modulate the field sent by the reader [84]. These tags are small and cheap; however, they restrict the reading range in between 2 mm to a few meters [84]. The semi-passive tags have an internal power source just to keep the chip powered up, but not for signal generation. Since it has internal power supply, the antenna is further optimized for backscattering to improve the reading range of the passive tags. Active tags are equipped with power supply that can energize the micro-chip and at the same time can generate signal to respond to the readers. Active tags’ reading range can be tens of meters.

In terms of their operating frequencies, RFID tags are categorized into three categories: low-frequency (LF), high-frequency (HF), and ultra high-frequency (UHF) tags. LF tags operate on 125–134.2 kHz and 140–148.5 kHz. They are cheaper than the higher frequency tags, but have the limitations of short reading range. HF tags operate at 13.56 MHz. They have higher reading ranges than the LF tags, but are costlier. LF and HF tags are license-exempt. UHF tags work on 868 MHz (Europe), 915 MHz (USA), 950 MHz (Japan), and 2.45 GHz. UHF operating frequencies vary from country to country and require a permit. The reading range of UHF tags varies from 3–6 meters for passive tags to 30 meters for active tags. These tags are costly and are used mostly for automated toll collection. RFID tags are used for a vast number of applications, ranging from inventory control [25, 26], animal tagging [64, 130], toll collection [76, 92], gasoline cards [18], money cards [15], and so on.
As amply proven by the Internet, standardization is essential for interconnection and interoperability of heterogeneous systems. The same holds for logistics, particularly in the context of shared 3PL where the provider must deal with a wide range of products, clients, requirements, and regulations. The recognition of success of the Internet has led to the paradigm called “Physical Internet” (PI), which is attempting to bring higher efficiencies to logistics by using the ideas of sharing, standardization, and automation [113, 122]. The PI initiative has an active consortium (see www.physicalinternetinitiative.org) and is being further propelled by advances in IoT technologies for rich, real-time sensing and consequent decision making and actuation. A key concept in physical internet is PI (or \( \pi \)) containers are defined to recursively compose to create bigger and bigger containers, and thereby lead to efficient recursive packing/unpacking. The GS1 set of International standards (see www.gs1.org) provide comprehensive RFID-based tagging and addressing of all important T&D entities including products, packages (including pallets), carriers (e.g., trucks, railcars, ships), warehouses, end-points, etc. Driven by the benefits in cost reduction and traceability, these standards are being rapidly adopted by the industry, and some retail chains like Whole Foods already mandate them.

The PI initiative considers logistics in the overall context of increasing efficiencies and reducing cost. In References [90, 119, 123, 126], we have considered PI from the perspective of distributing fresh foods where the quality deterioration and cost of spoiled product must be considered centrally, which is defined as Fresh Food Physical Internet (FFPI or \( F^2 \pi \)). We will further exploit these ideas in this article and put forward the vision of innovative sensing, communications, and analytics infrastructure to substantially improve the T&D fresh food supply chain.

3 FOOD QUALITY AND ENVIRONMENTAL FACTORS

Food products often deteriorate in quality or in value/usefulness as a function of flow time through the logistics system. The deterioration as a function of time \( t \) can be described by a non-decreasing function that we henceforth denote as \( \zeta(t) \). In general, \( \zeta(t) \) is linear for fruits or vegetables and exponential for fish/meat. The decay itself is a complex phenomenon and could refer to many aspects, including those that can be directly detected by the customers (e.g., color, texture, firmness, taste) and those that are latent but perhaps even more important, such as degradation of vitamin content or growth of bacteria. Furthermore, the decay rate is strongly influenced by the environmental parameters such as temperature, humidity, vibration, and so on.

**Temperature:** Temperature management is the most important and simplest way for delaying the quality deterioration of fresh foods, starting from production to distribution. In case of fresh fruits and vegetables, optimum quality is ensured only when they are promptly cooled to their optimum temperature after harvest and subsequently maintained under optimum temperature conditions. Several studies have shown that promptly pre-cooling berry fruits and subsequently maintaining optimum temperature significantly reduces loss of quality during storage and extends shelf life. Recent data show that if strawberries are not pre-cooled adequately after harvest, the waste at the retail level can be as high as 50% [59]. Also, fluctuating temperature during the distribution process often results in loss of shelf life.

**Humidity:** Humidity of the surrounding environment is also an important factor in food logistics, and should be maintained at a level that minimizes the water vapor pressure deficit between the product and the environment. When the relative humidity is too low, transpiration is enhanced, which results in loss of moisture. For example, strawberries stored at 10°C and 95% relative humidity experience ~7% more quality loss after four days than when stored at 75% relative humidity [147]. The use of protective packaging maintains a higher level of humidity and thus helps reduce loss of moisture.
Vibration: In-transit vibration results in quality degradation of the food products at the consumer end and results in reduced profit for the produce industry. Especially fruits like berries and grapes are highly susceptible to in-transit vibration, which causes skin abrasion, bruising, discoloration, and thus reduced shelf life. Authors in Reference [62] have studied that there is a range of frequencies that causes maximum damage to perishable products, which are found to be between 7.5 and 10 Hertz for these two commodities.

Controlled Atmosphere: Besides temperature and humidity, controlled atmosphere storage is also beneficial in reducing quality decay. For example, studies show that strawberry shelf life is increased by \(~4–8\) days when they are stored at \(5^\circ\text{C}\) in high \(O_2\) concentrations compared with storage in air [37]. Similarly, blackberries stored at \(2^\circ\text{C}\) in 10% \(O_2\) and 15% \(CO_2\) exhibits less decay compared to storage in air [164]. Such controlled storage not only results in less quality loss, but also gives more attractive appearance, texture, aroma, and nutritive value [33].

Several papers are devoted towards ambient sensing of food packages during transportation and storage. Authors in Reference [97] have proposed a vision of the “Intelligent Container,” which keeps track of the temperature history of the container carrying perishable products, and use this to estimate the quality they have at the present time. It also estimates how long the food quality will be all right and when the foods have to be disposed of, since their quality has fallen below a given threshold. In Reference [19] the authors have proposed a real-time perishable food supply chain monitoring system based on the ZigBee-standard wireless sensor network to sense different kinds of environment parameters related to food safety and quality, such as temperature, humidity, carbon dioxide, ethylene, vibration, and so on. Similar ambient sensing works are reported in References [38, 56, 101].

However, just ambient sensing of a few certain points in a container leads to wrong shelf-life prediction of the food packages. This is because the temperature of small packages inside a pallet vary depending on their positions inside the pallet. This is shown in Figure 7, where the temperature variation of 45 packages inside a pallet was recorded over a period of 24 hours [59]. As expected, the rate of temperature increase at the core of the pallet is lower than that of the outside. Thus, this results in a significant variation between the estimated remaining shelf lives of products. Figure 7(b) shows the distribution of estimated remaining shelf life of the 45 packages after 24 hours of storage. The minimum and maximum estimated shelf life are 410 and 470 days, while
Fig. 8. Impact of temperature on the appearance of "Killarney" red raspberries during a storage period of seven days (reprinted with permission from Reference [39]).

60% of the packages have a shelf life of less than 450 days. There is a wide variation of ∼60 days between the packages that are stored within the same pallet for 24 hours. This leads to the need of accurate quality sensing in every package level rather than sensing the temperature, humidity, and so on, within few specific points inside a container.

4 MODELING THE PERISHABILITY METRIC

Fresh food packages deteriorate in quality over time according to complex biochemical processes that depend on the food type, initial quality, temperature, humidity, vibrations, bacterial level, and bruises during storage/transportation. For example, Figure 8 shows the impact of storage temperature on the appearance of raspberries during a storage period of seven days. At any certain environmental conditions (temperature, humidity, etc.) the food quality deterioration as a function of time $t$ can be described by a non-decreasing function that is henceforth denoted as $\zeta(t)$. The deterioration of food products are nothing but biochemical reactions, thus $\zeta(t)$ can be modeled as a function of some measurable parameters related to the reaction that determines the quality loss.

4.1 Time-temperature Indicator and Arrhenious Equation

We describe a metric for estimating the degradation of a measurable parameter of a food product using its time-temperature indicator. Like any other biochemical reactions, in food products as well, certain parameter $A$ is converted to another, say $B$, over time. As the reaction proceeds, the concentration of $A$ decreases, whereas that of $B$ increases. The concentration loss (of $A$) or gain (of $B$) of a particular ingredient at any instance can be described by the following equation:

$$\frac{dC}{dt} = k.C^n,$$

where $C$ is the concentration of the ingredient, $k$ is the rate of degradation that depends on temperature and other factors, and $n$ is the order of reaction that is either 0 (zero order) or 1 (first order) for most food products. Food products such as fruits or vegetables generally follow zero-order degradation or linear decay, whereas products such as meat or fish follow first-order degradation of exponential decay. Thus, the quality can be measured by checking the concentration of certain ingredients in food products. For example, concentration of vitamin C or sulfur can be
quality indicators for different fruits or vegetables, whereas the concentration of bacteria such as Mesophiles, Psychrotrophs, Lactobacilli, Enterococci, Coliforms, and so on, can be the quality indicators for meat products.

The rate of concentration loss or gain \( k \) at different temperatures can be modeled using the Arrhenious equation as follows:

\[
k = k_0 e^{-\frac{E_a}{RT}},
\]

where \( k_0 \) is a constant, \( E_a \) is the activation energy, \( R \) and \( T \) are gas constant and absolute temperature, respectively. Thus, if a food product goes through multiple phases \( i = 1, 2, \ldots, m \) with time \( t_i \) at \( i \)th stage and rate \( k_i \) (based on temperature \( T_i \)), then the end concentration of an ingredient

\[
C = C_0 \sum_{i=1}^{m} k_i t_i = C_0 \sum_{i=1}^{m} k_0 t_i e^{-\frac{E_a}{RT_i}} \quad \text{for zero-order},
\]

\[
C = C_0 e^{\pm \sum_{i=1}^{m} k_0 t_i e^{-\frac{E_a}{RT_i}}} \quad \text{for first-order},
\]

where \( C_0 \) is the initial concentration of the ingredient. Notice that in Equation (3), \( t_i \) is calculated relative to the initial time \( t_0 \). The parameter \( k \) can also be obtained from experimental results at different temperatures. In Equation (3) the \( \pm \) sign represents the loss or gain of concentration depending on the type of the measurable parameter. For example, the vitamin concentrations of fruits or vegetables decay over time, whereas the bacterial growth on meat products adds up with time. Figure 9(a) shows the vitamin C degradation of different vegetables over time at 20°C, which shows linear decay, thus \( k \)'s are the slopes of the graphs. Figure 9(b) also shows similar characteristics of vitamin C decay over storage time for pineapples, capsicums, and green and red peppers. Also the exponential growth of certain bacteria on meat substances such as lamb meat and pork sausage are shown in Figures 9(c)–(d), where \( \ln(k) \)'s are the slopes.

With these, we can scale a perishability function of a food product \( \zeta(t) \) from the level of concentration of some measurable parameters. It suffices to assume that \( \zeta \) belongs to the real range \([0, 1]\) where 1 means that the product so far has not suffered any quality loss and 0 means that product has no value. Thus, if we know the storage time and temperature of the chain in different stages, we can estimate the delivery quality of a product using the above approach. In reality, because of different disturbances in the food chain (vibrations, change in temperature or humidity, etc.) the quality deteriorates faster or slower than the theoretical expected rate.

### 4.2 Shelf-life Computation

The kinetic modeling is also useful for measuring the shelf life \( t_{SL} \) of a product, which is the time taken by a product to reach the lowest acceptable quality (or concentration) \( C_1 \) given the initial concentration of \( C_0 \). Thus, the shelf life of a product can be calculated as

\[
\int_{C_0}^{C_1} \frac{dC}{C^n} = \int_0^{t_{SL}} k.dt \quad \Rightarrow \quad t_{SL} = \frac{\int_{C_0}^{C_1} \frac{dC}{C^n}}{k} = \frac{f(C_0, C_1)}{k},
\]

where \( f(C_0, C_1) \) is the quality function composed of the initial concentration \( C_0 \) and the concentration for maintaining a lowest acceptable quality \( C_1 \) of any ingredient in a food package. Thus, the quality function depends on the reaction kinetics that are governing the rate of decrease (or increase) of some chemical components or quality attributes.

Table 1 shows the reaction rate and the shelf life of different kinetic models. As mentioned in Section 4.1, most of the vegetables show zero-order kinetics, whereas the bacterial growth on meat substances follow first-order kinetics. In Reference [52], the authors have claimed that the oxidation of extractable color pigments in chili pepper follows second-order kinetics. Other than the
Fig. 9. (a) Vitamin C degradation in different vegetables at 20°C (data obtained from Reference [108]). (b) Changes in vitamin C content of green peppers, pineapples, capsicums, and red peppers during their storage at 1°C, 22°C, 5°C, and 28°C, respectively (data obtained from References [104, 133, 134, 160]). (c) Bacterial content in raw lamb meat kept for 0, 7, 14, and 21 days under retail conditions (data obtained from Reference [118]). (d) Observed bacteria growth on industrial pork sausage samples of two different trademarks at 4°C (data obtained from Reference [58]).

Table 1. Different Types of Kinetics Models

<table>
<thead>
<tr>
<th>Reaction type</th>
<th>Reaction rate</th>
<th>Shelf life</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zero-order</td>
<td>$- \frac{dC}{dt} = k$</td>
<td>$t_{SL} = \frac{C_0 - C_1}{k}$</td>
</tr>
<tr>
<td>First-order</td>
<td>$- \frac{dC}{dt} = kC$</td>
<td>$t_{SL} = \frac{\ln \left( \frac{C_0}{C_1} \right)}{k}$</td>
</tr>
<tr>
<td>Second-order</td>
<td>$- \frac{dC}{dt} = kC^2$</td>
<td>$t_{SL} = \frac{1}{k} \left( C_1 - C_0 \right)$</td>
</tr>
<tr>
<td>Logistic</td>
<td>$- \frac{dC}{dt} = kC \left( \frac{C_0}{C_{inf}} - 1 \right)$</td>
<td>$t_{SL} = \frac{\ln \left( \frac{C_0}{C_{inf}} \right)}{k}$</td>
</tr>
</tbody>
</table>

zero-, first-, and second-order reactions, another type of reaction that is studied in the literature is logistic model [103]. In logistic model, the rate of deterioration of a chemical content’s concentration is represented as $- \frac{dC}{dt} = kC \left( \frac{C_0}{C_{inf}} - 1 \right)$, where $C_{inf}$ corresponds to the concentration at the lowest quality. $C_{ba}$ is the constant representing the biological age of the product [157, 158].

Biochemical reactions are important to understand food quality monitoring, as many foods contain biological materials like enzymes. Sometimes these enzymes are desired (for instance, in cheese ripening) but mostly enzymes need to be deactivated to avoid deterioration of food quality; for example, enzymes like polyphenoloxidase lead to the browning of apples, potatoes, and cauliflower, and the formation of lipases gives raw milk a rancid taste, and so on. For modeling these biochemical enzymatic reactions, the Michaelis-Menten kinetics equation \[ 158 \] is used, which is written as

\[
\frac{dC}{dt} = \frac{V_{\text{max}}C}{C + K_M},
\]

where \( V_{\text{max}} \) is the maximum rate of reaction, and \( K_M \) is the Michaelis constant representing the enzyme concentration when the reaction rate reaches the half of the maximum rate [100]. Michaelis kinetics is used to study the relationship between the concentration of \( O_2 \) and the respiration rate of stored apples in Reference [67].

Physical processes such as creaming or sedimentation, fracture phenomena, viscosity changes, gelation of biopolymers, crystallization, and moisture migration also lead to food quality change. These phenomena are complex to model. An empirical relation of predicting the viscosity of dispersions can be derived from the Eilers equations as follows:

\[
\frac{\eta}{\eta_s} = \left(1 + \frac{1.25\phi}{1 - \frac{\phi}{\phi_{\text{max}}}}\right)^2,
\]

where \( \eta \) represents the viscosity of the dispersion, \( \eta_s \) the viscosity of the solvent, and \( \phi \) is the volume fraction of the dispersed particles. This model is applied to model the skimmed milk samples with varying volume fraction of casein micelles in Reference [34]. In Reference [105], the authors have used this model to describe the influence of dispersed particles on deformation properties of concentrated caseinate composites. Different types of physical models can be found in References [162, 165].

Microbiological changes in fresh food happen due to the growth of different microorganisms. Sometimes these growths are desirable (such as in fermentation), but mostly microbial growth is undesirable, as it leads to spoilage and thus illness in the human body. Thus, predicting microbial growth like bacteria in foods is of utmost importance for predicting shelf life. Microbial growth can be modeled by using the Gompertz model, which can be described as follows:

\[
\ln \frac{N}{N_0} = A_s \exp \left[-\exp \left(\frac{\mu_{\text{max}} e}{A_s} (\lambda - t) + 1\right)\right],
\]

where \( N \) is the number of microorganisms, \( N_0 \) is the number of microorganisms at time zero, \( A_s \) is the asymptotic value of the maximum number of microorganisms, \( \mu_{\text{max}} \) the maximum growth rate, and \( \lambda \) is the lag phase. The Gompertz model is applied to study the growth of Salmonellae in Reference [162]. Some other microbial growth models in fresh food are reported in References [8, 109].

5 FOOD SENSING TECHNOLOGIES

Food sensors embedded in intelligent packaging provide the sellers as well as the buyers reliable information on the food, how carefully it is handled during the transportation and storage, as well as the packaging integrity. The simplest form of intelligent packaging is the use of smart labels that display the shelf life of the food packages. For example, a time temperature indicator (TTI) is a smart label that shows the accumulated time-temperature history of a product.

There are several TTI technologies available on the market. TTI can be placed as a sticker on transport containers. If a food container is exposed to different recommended temperatures, then
an irreversible chemical change will be reflected on the indicators [94]. TTI is mainly useful to check the quality of chilled or frozen food, where cold storage is a critical control point during the transport and distribution to indicate probable food degradation.

Various chemical TTI systems have been developed in recent years, which are primarily polymerization-based, photochromatic-based, and oxidation-based [166]. Fresh-Check® [17] and Heatmarker® [23] are the examples of polymerization-based TTI. Fresh-Check is composed of two concentric circles, as shown in Figure 10(a). A polymer is set in the inner ring, which gradually changes color as the temperature and reaction time increases. By assessing the color of the inner ring in contrast to the outer color, the quality of the product is analyzed. When the product reaches the end of its shelf life, the inner color becomes darker than the outer one. Similar other polymerization-based TTIs are reported in References [40, 131]. Photochromatic-based TTIs are activated by certain wavelengths of light. One such chemical TTI is OnVuTM [13], which is activated by UV light. OnVu consists of a temperature-sensitive intelligent ink at the “core” and a reference color at the “shell,” designed in the shape of an apple. With the change in temperature, the color of the photochromatic material at the core fades and become colorless. The oxidation-based TTIs react with oxygen and shows color change to monitor the accumulation of time and temperature [68, 166].
Other types of TTIs are diffusion-based, enzymatic, or biological. Monitor Mark\textsuperscript{TM} [7] and Tempix [11] are examples of diffusion-based TTIs that are based on temperature-dependent diffusion reactions of certain substances. FreshTag [4], developed by Vitsab, is an example of enzymatic TTI, which is green when activated, then turns yellow, and finally red when it reaches its endpoint. FreshTag TTI is illustrated in Figure 10(b). In case of biological TTIs, yeasts, and bacteria like lactic acid bacteria, Streptococcus are used to measure the change in pH that leads to color change [12, 98, 163].

However, indicator sensors are used to analyze the quality of the food packages by monitoring different organic compounds, ethanol, glucose, or gas molecules such as $H_2$, $CO_2$, $H_2S$, $NH_3$, $CH_4$, and so on [148]. Such indicators are useful to understand the toxic composition of the gases produced from decomposing food inside a food container [107]. These changes are detected by the indicators and transformed into a response, usually by showing a color change on the smart labels. Methyl red is another example of an indicator sensor that works on the increase in pH, due to volatile amines decomposition and is used for real-time monitoring of fresh chicken [96]. RipeSense is another first intelligent sensor label that changes color to indicate the ripeness of the fruit [135], which works through the reaction of the aromas released by the fruit as it ripens. As shown in Figure 10(c) RipeSense sensors are initially red, then gradually turn orange, and finally yellow. A similar sticker that indicates whether a fruit or vegetable is ripe or not is RediRipe [9], which detects ethylene gas, which is released by fruits/vegetables as they ripen. Using such sensors, food suppliers can understand the ripeness of fruits stored in the warehouse and determine the best time to deliver them to the supermarket. DSM NV has developed SensorQ to sense and report the formation of biogenic amines in microbiological origin meat and fish [55]. Authors in Reference [114] have also developed a new colorimetric sensor for monitoring the deterioration of fish.

Other than food quality sensors, developing food contamination detection sensors is currently a very active field as well, with sensors ranging from simple to highly sophisticated, with functionalities including bacterial content, contamination, texture/color degradation, bruising, and so on. Simple sensors are beginning to show up in packages sold to end customers, such as $C_2$Sense [20], FoodScan [146], Salmonella Sensing System [14], and so on. $C_2$Sense sensors detect ethylene, biogenic amines, and other relevant gases at low ppm and sub-ppm concentrations. They can monitor fruit ripeness and meat/fish/poultry freshness at all stages of the supply chain. FoodScan can detect the presence or absence of bacteria and contaminants in food before anyone can see or smell them. The Salmonella Sensing System invented by Auburn University uses a magnetoeelastic biosensor that is coated with a bacteria-specific recognition layer capable of detecting specific types of pathogenic bacteria. Michigan State University [172] has invented a biosensor that can detect Escherichia coli O157:H7 and Salmonella in meat products. A biosensor developed by the Naval Research Laboratory [145] can detect Staphylococcal enterotoxin B and Botulinum toxin A in tomatoes, sweet corn, beans, mushrooms, and tuna. Molecular Circuitry Inc. [5] has developed a biosensor that can detect Escherichia coli 0157, Salmonella, Listeria, and Campylobacter. Many such food sensors are summarized in Table 2.

Such tiny quality sensors can be embedded in food containers for use as electronic bar codes to monitor food in all its phases from production to processing, distribution, and consumption. However, such bar codes are investigated at only a few key points of the food chain. Thus, the deterioration in quality cannot be generally identified until it becomes too late to take any proactive actions for avoiding such quality loss. For such continuous monitoring purposes, such tiny sensors should communicate with each other and dynamically self-organize necessarily to build an online sensing infrastructure. Therefore, an important line of investigation is to consider future scenarios where not just the sensor, but the entire sensor module (including the communication module) is inexpensive enough to be embedded in each package within a box.
### Table 2. Different Sensor Types, Data Obtained from References [57, 66]

<table>
<thead>
<tr>
<th>Types</th>
<th>Trade Name/Biosensor</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTI</td>
<td>Fresh-Check</td>
<td>Temptime Corp.</td>
</tr>
<tr>
<td></td>
<td>Timestrip</td>
<td>Timestrip Plc</td>
</tr>
<tr>
<td></td>
<td>OnVu</td>
<td>Ciba Speciality Chemical and Freshpoint</td>
</tr>
<tr>
<td></td>
<td>MonitorMark</td>
<td>3M™, Minnesota</td>
</tr>
<tr>
<td></td>
<td>Tempix</td>
<td>TEMPIX</td>
</tr>
<tr>
<td></td>
<td>FreshTag</td>
<td>Vitsab</td>
</tr>
<tr>
<td>Indicator Sensors</td>
<td>SensorQ</td>
<td>DSM NV</td>
</tr>
<tr>
<td></td>
<td>RipeSense</td>
<td>RipSense™</td>
</tr>
<tr>
<td></td>
<td>RediRipe</td>
<td>University of Arizona</td>
</tr>
<tr>
<td>Contamination</td>
<td>C₂Sense</td>
<td>C₂Sense, Inc.</td>
</tr>
<tr>
<td>Sensors</td>
<td>FoodScan</td>
<td>MS Tech</td>
</tr>
<tr>
<td></td>
<td>Salmonella Sensing System</td>
<td>Auburn University</td>
</tr>
<tr>
<td></td>
<td><em>Escherichia coli</em> 0157:H7 in lettuce*</td>
<td>MIT</td>
</tr>
<tr>
<td></td>
<td><em>Escherichia coli</em> O157:H7 and Salmonella in meat</td>
<td>Michigan State University</td>
</tr>
<tr>
<td></td>
<td>Salmonella and Campylobacter in pork industry</td>
<td>Georgia Research Tech Institute</td>
</tr>
<tr>
<td></td>
<td><em>Staphylococcal enterotoxin B and Botulinum toxin A in tomatoes, sweet corn, beans, and mushrooms</em></td>
<td>Naval Research Laboratory</td>
</tr>
<tr>
<td></td>
<td><em>Escherichia coli</em> 0157, Salmonella, Listeria, and Campylobacter*</td>
<td>Molecular Circuitry Inc.</td>
</tr>
</tbody>
</table>

### 6 COMMUNICATING SENSED DATA

An integral component of the sensor module designed for automated operation is a radio that can receive configuration commands, send out sensed data, and forward received data to support multihop communications. In the fresh food environment, communications should be possible across radios in different boxes in a pallet; therefore, the range requirements are very short, but the signals should be able to propagate well through aqueous tissue media. Unfortunately, the normal RF communication (e.g., Bluetooth in the 2.4 GHz ISM band) does not work very well in this environment due to high signal absorption and complex channel conditions [85]. Reducing absorption by choosing lower frequencies (e.g., 802.11 ah) helps in attenuation, but would need bigger antennas and cause severe interference in this very dense sensor environment. Authors in Reference [60] show that in body-area-networks, especially in tissue medium, the path loss at 50 mm is 47–49 dB at 403.5 MHz. Other studies [51, 167] have also reported attenuation values ranging from 20 dB at 100 MHz to 60 dB at 1 GHz for distances less than 10 cm. High RF transmission powers are certainly undesirable because of serious overheating and thus quality loss [144].

Another promising technology in food or tissue medium is acoustic or ultrasonic communication. Ultrasound transducers operate at frequencies above the human hearing range [140]. Ultrasonic waves are generated by piezoelectric materials that convert electrical energy to mechanical energy and vice versa [30, 93, 149]. Ultrasonic waves are subject to lower absorption as compared to RF waves in aqueous or tissue medium. Because of these characteristics, ultrasound waves are used successfully for underwater communications [32, 110], wireless body area networks [70, 141, 143],
and ultrasonic imaging [136, 155]. In Reference [80], the authors studied ultrasonic communication between nanorobots across various distances. Authors in Reference [142] successfully built an ultrasonic testbed based on USRP2/N210 software-defined radios interfaced with low-frequency (LFRX and LFTX) daughterboards connected to ultrasonic transducers. They have used medical training phantoms (human kidney) in between transmitters and receivers to emulate propagation through tissues. The attenuation reported is 2 dB/cm at 5 MHz.

However, ultrasound (or acoustic) propagation in tissues is deeply affected by multipath fading because of the inhomogeneity of the medium in terms of density, presence of very small organs and particles, and so on, which leads to the variation in sound velocity. Due to the presence of such reflectors and scatterers, numerous attenuated, distorted, and delayed versions of the same transmitted signal reach the receiver, making detection and decoding a challenging operation. Moreover, the low speed of sound in water or tissues leads to high propagation delay, which makes the medium access schemes problematic in a highly dense environment. Also, a significant portion of the energy is absorbed (although lower than RF) and converted into heat when ultrasounds propagate. This could potentially lead to a temperature increase, which will affect the quality of perishable food being transported.

Such challenges cannot be overcome until a major paradigm shift in communications through food medium is made to address the limitations of current typical communication technologies. Prior research indicates that the Magnetic Induction (MI)-based communication [74, 127] in the HF band (3–30 MHz) is largely unaffected by the tissue medium [106]. MI communication is based on the principle of resonant inductive coupling (RIC), which involves two matched coils, each forming an LC circuit with the same resonance frequency, as shown in Figure 11. RIC has been used successfully for extremely efficient power transfer over short distances and is used for contactless mobile charging, car battery charging, and so on.

MI communication modulates the magnetic field and forms the basis of near field communications (NFC) between mobile devices. Compared to RF, MI suffers from smaller signal fluctuations and multi-path effects. The ability to use small coils (2.5/5.0 cm in our experiments), short transmission range (e.g., 1.5 m) and a decent data rate (e.g., 596 Kb/s) are ideal for our application. Finally, ultra-low-power MI solutions consume less power than RF and ultrasound over short distances. The MI channels are also more determined, as the MI signals are not reflected or scattered by the surrounding environments, and thus suffer from smaller signal fluctuations and multi-path effects. Also, there is no known biological effect of MI in the food or tissue medium [127]. A comparison between three modes of communication is depicted in Table 3.

We thus consider MI communication as the most viable communication framework in this context [124, 125] and describe some preliminary analysis on the feasibility of such communication. The magnetic coupling between two coils depends on their relative orientation, and the energy transfer goes down as cosine of the relative angle between the coils. The magnetic field generated
Table 3. Comparison between RF, Acoustic, and MI

<table>
<thead>
<tr>
<th></th>
<th>Absorption</th>
<th>Multipath</th>
<th>Datarate</th>
<th>Power</th>
<th>Size</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>High</td>
<td>Moderate</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Ultrasonic</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>MI</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

by a unidirectional coil can be characterized using Lenz’s law of electromagnetism [39]. Figure 12(a) shows the path-loss in between two such coils that are one meter apart. The parameters $R_{\parallel}$ and $R_{\perp}$ specify the rotation of receiver coil relative to the transmitter along the same and orthogonal axes, respectively. For this plot, we assume 10 turns for each coil and an operating frequency of 10 MHz. It is seen that the loss is minimal when the coils are aligned and maximum when the coils are orthogonal.

Such variability in coupling would require the transmit and receive coils to be well aligned, which may be difficult in practice. However, the issue can be resolved by using an antenna consisting of three concentric coils, one in each plane ($xy$, $yz$, $xz$). Figure 12(b) shows the path-loss in this case (all other parameters same as for Figure 12(a)). It is seen that the signal is almost isotropic in this case, which means that an S&C module hosting such antennas can be packed into the boxes in any orientation. Notice that in a tri-directional coil, the orthogonal coils on the same wireless device do not interfere with each other, since the magnetic flux generated by one coil becomes zero at the other two coils.

We next use two tri-directional coil transceivers and vary the distances in between them to measure the transmission range of the magnetic communication. Figure 13 shows that the transmission range is typically 2–3 meters (considering 70 dB of loss) when the coil radius ($\rho$) = 2.5 cm and even higher with a larger coil. This is adequate for our application.

We have used 10 MHz frequency in the illustration based on its use in the previous research literature [31, 106, 153]. At 10 MHz, the wavelength is 30 meters, which means an NFC range of 4.8 meters and is adequate for our applications. We have conducted an experimental evaluation of an experimental MI radio from Freelinc [1] in different media and have achieved a transmission range of 1.5–3 meters at 13.56 MHz [120]. However, the channel conditions at a wide range of frequency levels (from hundreds of kilohertz to few tens of megahertz) must be evaluated through a large variety of fresh food media, including fruits/vegetables with different water/mineral/salt contents and meat with different tissue characteristics. Furthermore, while tiny coils (less than a
Fig. 13. Path loss with distance ($r$) for tri-directional coil transceivers.

Fig. 14. Sensor in a box.

centimeter in diameter) are desirable, they are likely to pose additional challenges because of coil thickness being a significant percentage of coil diameter. In fact, how the coils should be wound and how the coil pair should be tuned to provide most efficient magnetic coupling become very challenging, particularly when tri-directional coils are considered. In addition to that, building low-power MI communication protocols in a dense environment is quiet sparse, which needs more research. In our recent work [120], we have explored some low-power MI communication protocols through silence, along with the use of multiple MI channels to reduce the effects of network overhearing.

7 SENSOR NETWORK FOR MONITORING FRESH FOOD

To do a comprehensive online monitoring of the fresh food quality, it is necessary to implement a local sensing and communications infrastructure for a closely spaced set of pallets that may be carried on a truck, railcar, boat compartment, and so on, or simply sit in a cold storage warehouse room. Each pallet consists of some number (few 10’s to few 100’s) of retail boxes, which in turn may either contain the fresh produce directly or the customer-level product packages (e.g., for meat). Figure 14 shows a sensor attached to a chicken package in a box of such packages. We assume that each retail box contains a small battery-operated sensing and communication (S&C)
module consisting of some sensors and a radio. The sensors could be contact type (e.g., chemical) or non-contact (e.g., gas sampling, imaging). We assume that all sensors are removed and returned at the end-point (e.g., at a retailer) to the shipper/processing plant via the normal reverse logistics that has to be in place to handle distribution of products returned/rejected by the destination. The sensors could be checked and then reused. This reuse is essential at least for the foreseeable future while the sensor cost remains significant. In the long run, it may be possible to simply embed the sensors in the boxes, which may or may not be reused.

We assume that each box has GS1 compatible RFID tag (already in practice) that is tracked during loading/unloading (manually or mechanically). We assume the box order and their RFID is recorded during palletization so the relative box location vs. RFID map can be generated automatically. A truck/warehouse may carry many such pallets, and we assume that pallet RFID and their relative position in the truck/warehouse is also recorded (manually or mechanically).

The radio in the S&C module provides two functionalities: (i) depositing sensed data with a sink or anchor node and (ii) localizing the containing box so it is possible to identify the box(es) having quality issues. We assume that the sensor module is packed in a nearly identical spot within each box so it is possible to localize a box (or rather the sensor radio of the box). The localization can tell us which relative box in a pallet (or truck) has a product with quality/contamination issues. Given the loading order vs. RFID map discussed above, we could thus correlate the position of this box with the box and pallet RFIDs and thus enable RFID-based tracking.

It is important to note that the regularity assumptions made above are becoming increasingly realistic due to the rapid automation of logistics, including those related to Industry 4.0 and Physical Internet initiatives. In other words, we expect the mechanisms discussed in this article to be increasingly applicable and even essential as the full potential of Physical Internet and Industry 4.0 is realized.

Figure 15(a) shows the overall architecture for the proposed system by considering deployment within a carrier (a truck) carrying one or more pallets of boxes with one sensor per box. The sink (or anchor) nodes are mounted on (or close to) the inner walls of the carrier. They can communicate with sensor modules and among themselves via RF, acoustic, or through magnetic induction (MI). We also integrate a Wi-Fi interface in some of the anchors for communications outside the span of product boxes. The Wi-Fi-enabled anchors can connect to a Wi-Fi access point, which is shown in the figure as the driver’s smartphone but could also be a more permanent and better secured device installed in the carrier. Similar deployment happens in a cold-storage warehouse room as well, which is shown in Figure 15(b).

The architecture is designed to relay all sensed data to the dual-interface anchors and then to the local hub, which in turn sends all data to an Analytics & Operations Center (AOC) via a cellular

---

Fig. 15. Communication Infrastructure (a) inside a truck and (b) within a warehouse.
link for comprehensive analytics and control. The AOC could be hosted in a cloud and could be physically (but not logically) distributed, but that aspect is not important here. The AOC receives data periodically from all active carriers and warehouses (including empty ones) operated by the 3PL operator in its entire network, so it has a global view of the operations and resources. The centralized availability of sensed quality, contamination, and localization information at the AOC from the entire T&D network has the potential to revolutionize logistics, particularly when coupled with ongoing automation trends.

8 DATA ANALYTICS FOR PROACTIVE DECISION MAKING

As shown in Figure 15, each box will contain the S&C module, which not only results in a very dense radio environment but also requires multi-hop communications to send the food quality data from any arbitrary box to the local hub. Since our data collection frequency is very low (e.g., quality sensed and transmitted every 30 minutes), we can keep all S&C modules mostly in a deep sleep mode. At the next data collection, they can wake up according to a fixed schedule, collect and transmit its data (along with its relative location) to a neighbor and go back to sleep. By starting with the innermost modules in the pallet and moving towards outer ones that are closer to the anchor nodes, it is easy to devise a simple store and forward mechanism to push data from each S&C module to an anchor via MI communications, and from there to the local hub over Wi-Fi. Our simple calculations suggest that it is possible to run an S&C module for more than a year using inexpensive button batteries.

Assuming a ubiquitous deployment of the infrastructure in Figure 15 in the T&D system, the AOC can collect a fine-grain view of all of the T&D operations including the status, location, and condition of every product in the T&D pipeline. This allows for online analytics of collected data from individual carriers and warehouses, system-wide analysis and optimization across the entire company, and derivation of insights and fine-tuning of the operations based on the offline analysis of historical data. In the following, we discuss three significant applications enabled by these capabilities that can substantially reduce cost and food wastage and increase efficiency. We also discuss how the analytics capabilities can be exploited for accruing nearly all of the advantages with only a sparse deployment.

8.1 Intelligent Distribution

Product quality and freshness are the key drivers of why buyers choose a brand or store [6]; thus, maintaining the delivery freshness is of primary importance to the grocers. Because of this competitive pressure for an attractive presentation, typically grocers incur a loss of 10% of what they purchase [75], whereas rejected shipping due to quality loss typically incurs a loss of 5%–10% revenue to the grower [132]. Total food losses in the USA are estimated at ~33% of that of harvest [21]. A key strategy in reducing food waste and simultaneously serving the societal needs is to distribute food proactively instead of rigidly transporting it from intended source to the destination. There are two key challenges in a proactive distribution: (a) technological and (b) socioeconomic. The technological challenge lies in online collection of product quality information and other processes in place to redirect carriers to do proactive distribution while still satisfying all other objectives. We believe that FFPI can go a long way in enabling this capability. The socioeconomic challenge relates to the parties involved willing to and being comfortable with information sharing, flexible contracts, and somewhat variable product quality.

Implementing such approaches can cut the losses in half for the growers and retailers [21]. Below, we describe some of the well-known distribution strategies along with their pros and cons.

First In First Out (FIFO): Typically the warehouses and distribution centers use a First In First Out (FIFO) (Figure 16(a)) distribution policy to determine which pallets to ship next, which is based
First Expired First Out (FEFO): A smarter and more efficient approach is to enable a First Expired First Out (FEFO) [79] (Figure 16(b)) approach by using the accurate shelf-life estimation from the online sensing infrastructure. From this figure, we can observe that with FIFO strategy, the first pallet is shipped first, because of the assumption that the pallet has arrived first and thus has the least shelf life. However, with accurate shelf-life information, that pallet is shipped third in case of FEFO, as two other pallets have lesser shelf life even if they have arrived later. This gives pallets with lower shelf life the opportunity to be consumed sooner than the other ones, which reduces the chance of wastage.

Intelligent distribution: Much better distribution strategy can be adopted by matching the remaining shelf life of each package to the transit time and consumption rate of each destination, as shown in Figure 16(c). In this strategy, pallets with low shelf life will be transported to the closest destinations but have faster consumption rate. However, pallets with longer shelf life will be transported to the destinations that are far apart and/or have lower consumption rate. Such intelligent distribution can be modeled using the following optimization problem:

\[
\begin{align*}
\text{Maximize} & \quad \sum_{i=1}^{n} \sum_{j=1}^{m} U(S_i - T_j - C_j) x_{ij} \\
\text{subject to} & \quad \sum_{j=1}^{m} x_{ij} = 1, \forall i,
\end{align*}
\]

where \( n \) and \( m \) are the number of pallets and destinations, respectively, \( x_{ij} \) is a binary decision variable that is 1 if the \( i \)th pallet is transported to the \( j \)th destination and zero otherwise. \( S_i \) is the shelf-life of the \( i \)th pallet, whereas \( C_j \) is the consumption time of the \( j \)th destination based on their local demands. \( T_j \) is the transit time to the \( j \)th destination. \( U(y) \) is the utility function that is 1 if \( y \geq 0 \) and 0 otherwise. In the optimization problem (8), the objective is to maximize the overall 

Fig. 16. (a) FIFO vs (b) FEFO vs (c) Intelligent distribution. The figure is adapted with permission from Chain-Link Research (www.clresearch.com) [21].
utility, which is the number of pallets that have non-negative shelf life at the time of consumption. The constraint states that all the pallets are transported to a destination.

A comparison of FEFO vs intelligent distribution is shown in Figures 16(b)–(c). In case of FEFO, pallets 2 and 4 are sent to destination 1, as they have the minimum shelf life and destination 1 has ordered before destination 2. However, this strategy assumes that the order that is placed earlier (destination 1 in this case) will be reached faster and the pallets will be consumed faster. However, in reality, destination 1 may take more transit time and have slower consumption rate, which may lead to quality loss and spoilage to pallets 2 and 4. For example, given the transit time and consumption rate of Figure 16(c), pallets 2 and 4 will be expired before consumption at destination 1. Using the objective function of Equation (8), the utility of FEFO distribution will be 2 as only pallets 3 and 4 will be consumed at destination 2 before their expiration. By solving the more informed intelligent distribution strategy using the optimization problem (8), we can observe that to maximize the overall utility, pallets 2 and 4 should be sent to destination 2 for faster consumption, whereas other pallets can be transported to destination 1, as they have higher shelf life. The overall utility in this distribution strategy is 4, as all the pallets are consumed before their expiration.

The global view provided by our architecture enables the operator to design suitable distribution strategies; however, they are by no means routine. The problem is that there are several practical restrictions on the extent to which a scheduled delivery of product X in the amount N from source S to destination D can be altered. These include several coupled factors: (a) willingness of the end-points to absorb excess or altered product as a function of price and storage capacity, (b) impact of type (i.e., perishability characteristics), quantity, and condition of products involved, and (c) understanding to what extent a change in schedule is possible and its consequences. However, the trucks may also need to change their delivery plans for unexpected congestion, road blocks, or because of the unexpected quality deterioration of the carrying packages. This needs tight interaction in between the shipping trucks along with the 3PL controller to ensure fresh and fuel-efficient delivery.

The related issues are of lateral distribution, e.g., local balancing of inventory across regional distribution centers or even retail stores, and distribution of food that is in danger of being spoiled soon to food kitchens or cooperatives at a very low cost. It is likely that a better designed and operated food network will address even more of the waste via quick local distribution of imperfect items and by delivering fresher and more palatable food to customers. Thus, in a flexible food network model, the operators can distribute the food products from the excess points to the shortage points within a neighborhood area. The shortage points may be the retailers that have sudden demands, or can be the food banks where they are consumed by people at a lesser cost. The excess/shortage points update their corresponding supplies/demands to their local 3PL operators, which take and inform the distribution decisions for the retailers.

The lateral transfer is also connected to the continuous sensing capability of the container sensors. For example, if a long-distance truck on its way finds some food packages deteriorating unexpectedly, then they can be distributed to nearby stores or food banks rather than carried all the way to their destined location and by that time are spoiled. Such lateral transfers can also happen from retailers to nearby retailers with a sudden shortage of supply or to nearby food banks. This lateral transfer is also connected with product substitution. Retailers order fresh vegetables and packets of specific types, brands, colors, or sizes based on their demand but in case of a supply shortage or unexpected quality deterioration, some products can be substituted with similar types and/or of different brands. As an example, an order of romaine blend with carrots from a manufacturer can be substituted by red leaf lettuce from a manufacturer [61]. However, there is an extra cost associated with the transportation in this lateral transfer process. For example, distributing
Table 4. Food-borne Illness Outbreaks in Recent US History [24]

<table>
<thead>
<tr>
<th>Outbreak Types</th>
<th>Details</th>
<th>Casualties</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salmonella outbreaks</td>
<td>Peanut Corporation of America (PCA) peanut butter</td>
<td>714 sickened, 9 died, 3,600 products recalled, PCA is now bankrupt.</td>
<td>2009</td>
</tr>
<tr>
<td>Cargill ground turkey</td>
<td>136 illnesses across 34 states, 1 died, 36M pounds of ground turkey recalled.</td>
<td></td>
<td>2011</td>
</tr>
<tr>
<td>Foster Farms chicken</td>
<td>634 illnesses across 29 states, voluntary recall on all Foster Farms brand chicken products.</td>
<td></td>
<td>2013</td>
</tr>
<tr>
<td>Mexican cucumbers</td>
<td>907 illnesses across 40 states, 6 deaths, Andrew &amp; Williamson Fresh Produce issued 2 separate recalls.</td>
<td></td>
<td>2015</td>
</tr>
<tr>
<td>Escherichia coli outbreaks</td>
<td>Dole baby spinach</td>
<td>31 suffered kidney failure, 205 reported cases of diarrhea and dehydration, 3 died, Dole recalled all its bagged spinach across the country.</td>
<td>2006</td>
</tr>
<tr>
<td>Taco Bell fast food</td>
<td>Eight people developed kidney failure, 53 were hospitalized.</td>
<td></td>
<td>2006</td>
</tr>
<tr>
<td>Chipotle Mexican Grill fast food</td>
<td>55 people in 11 states became ill.</td>
<td></td>
<td>2015</td>
</tr>
<tr>
<td>Botulism outbreaks</td>
<td>Trini &amp; Carmen’s hot sauce</td>
<td>58 people became ill, restaurant closed.</td>
<td>1977</td>
</tr>
<tr>
<td>Home-canned potatoes</td>
<td>29 people became ill, 1 death.</td>
<td></td>
<td>2015</td>
</tr>
<tr>
<td>Listeria infections</td>
<td>Hot dogs</td>
<td>Affected at least 100 people across 24 states, 14 adult deaths, and four miscarriages.</td>
<td>1998-1999</td>
</tr>
<tr>
<td>Pilgrim’s Pride turkey meat</td>
<td>7 adult deaths and 3 stillbirths, 27.4 million pounds of poultry products recalled.</td>
<td></td>
<td>2002</td>
</tr>
<tr>
<td>Cantaloupes</td>
<td>147 illnesses, 33 deaths.</td>
<td></td>
<td>2011</td>
</tr>
<tr>
<td>Hepatitis A infections</td>
<td>Frozen strawberries</td>
<td>affected 153 people.</td>
<td>1997</td>
</tr>
<tr>
<td>Chi-Chi’s salsa and chili con queso</td>
<td>155 illnesses, 3 deaths.</td>
<td></td>
<td>2003</td>
</tr>
<tr>
<td>Tropical Smoothie Cafe drinks</td>
<td>143 illnesses, 56 hospitalized.</td>
<td></td>
<td>2016</td>
</tr>
</tbody>
</table>

some deteriorating packages to local food banks at cheap prices may be beneficial rather than sending them to a faraway retailer from a business point of view. Trading off these competing design objectives is thus one of the challenges of this integrated infrastructure.

Notice that the success of intelligent and informed distribution depends on to what extent the shelf life of the individual packages can be estimated from the available sensing values. As there are multiple and dynamic sensing parameters and quality attributes associated with the food packages, shelf-life prediction with a good level of accuracy is a non-trivial problem. Various machine learning models (such as neural networks, fuzzy logic, evolutionary algorithms, genetic programming, etc.) can be applied for predicting the shelf life of products in this scenario. Such models have been applied for prediction of the shelf life of meat packages [35], processed cheese [73], and fruits and vegetables such as pomegranates [139], cabbages [171], sweet cherries [154], and so on. Reviews of different machine learning models in food sciences can be found in References [72, 83].
8.2 Proactive Contamination Detection

The publicly accessible extensive CDC database [47] shows only a modest decrease in FBI outbreaks and no decrease in hospitalization/deaths during 1998–2015 in spite of the 2008 food safety act [161]. Table 4 shows different types of food-borne illness and the amount of casualties caused by the outbreaks in the USA. This can be attributed to increasing consumption of fresh/untreated foods and a fragmented and largely manual food safety regime [91]. As a concrete example, the 2006 E. coli outbreak for Dole prepackaged spinach started in early July and after an intensive investigation, the farm of origin was confirmed only on October 5, 2006 [10]. Figure 17 shows the progress of the investigation and the number of cases reported during the investigation period, which shows the necessity of quicker/automated contamination detection.

Undoubtedly, the problem of automated tracing of source of contamination/FBI is extremely complex; however, we believe that our architecture in Figure 15 coupled with ongoing GS1 standard implementation and logistics automation can provide two key benefits: (a) proactive monitoring of contamination and associated removal of contaminated food from the supply chain and (b) ability to narrow down the potential sources of reported FBI/contamination. A difficult part of (b) is the ability to track purchase/consumption of product by individuals, since our end-points are retailers and other businesses. However, during the course of investigation, the FBI reports can be linked to the contaminated product RFIDs in other ways.

For traceback, the AOC can build a Bayesian inference–based graph, where the nodes represent all uniquely addressable entities and arcs represent the associations, i.e., a package (with certain RFID) carried via specific carriers, passing through specific warehouses and processing plants, sourced from a specific farm, and so on. Given such a graph and contamination reports, the traceback is conceptually straightforward but practically difficult for the following reasons:

(1) Building the entire graph explicitly and keeping it up to date is impractical in a large system, particularly to tackle relatively infrequent events of systematic contamination.

(2) Not all relationships are known with certainty; for example, two different products processed in the same plant may or may not be handled by the same machines or people.
(3) Given that sophisticated sensors are used sparingly for cost reasons, we do not necessarily know about the contamination status of every box.

(4) The observed contamination incidents may be a result of multiple unrelated but concurrent root causes.

One potential approach is to take a “bird’s eye view,” or coarse resolution, to estimate the possible origin points based on the sensory outcomes, and then search at progressively finer resolutions as necessary (if further information is available). To tackle the first problem, a compressed representation of keeping track of each and every packet’s provenance information is needed. Such compressed representations for keeping provenance information has been studied extensively in the context of web graph compression [50, 169, 170]. For tackling the other points (2–4), a macroscopic approach can be adopted, where the likelihood that the contamination origin lies within a certain region is first identified. For example, in case of the bacterial outbreak of 2006, few farms from Monterey County, CA, were identified as the contamination origin, whereas the outbreak affected over 200 people from 26 states. Thus, the first job is to identify the region of origin of such outbreaks from the inference model. This inference model next has to be further refined as more and more information or surveys are available within the entities of the pipeline.

8.3 Ambient Control

Refrigeration of fresh food to maintain quality and safety is very energy- and carbon-footprint intensive, particularly when used on moving carriers where it is dependent on gasoline/diesel. We believe that an intelligent management of climate control (temperature, humidity, airflow, etc.) can reduce this footprint substantially, particularly when coupled with modular cooling technologies for partially full carriers. Figure 18 illustrates the relationship between external factors, their impact on intrinsic properties of the foods, and the resulting impact on food quality. The problem of applying optimal cooling is extremely complex, because of complex dependence on numerous factors such as product type/condition, ambient conditions, cooling system characteristics and losses, total transit time, and so on. It becomes even more complex when a carrier carries multiple product types with different degradation processes, loading/unloading points, and freshness requirements. The T&D must also satisfy governmental (e.g., USDA) regulations [161] and cope with uncertainties in transit/loading/unloading delays, weather, starting condition, and so on.

In a typical food chain, environmental monitoring is done at few key points. Thus, the environmental variations in each pallet level is not captured. However, there are significant temperature variations at each pallet in a lot, which results in shelf-life variations between the pallets. Thus, the pallet-level monitoring of the integrated infrastructure enables a better opportunity for ambient
management. The freshness quality of the boxes also varies because of their handling behavior. For example, it takes roughly 15 minutes \cite{21} to pick and build a pallet of lettuce in a farmland. Thus, it takes several hours to pack a truck full of lettuce pallets, which may result in shelf-life variations of around three days \cite{138} between the first and last pallets. In fact, harvested crops sometimes remain in the field even longer, which leads to a larger variation in shelf life. In the transit stage, non-uniform ambient temperature in the trailers can lead to roughly 12°C of temperature variation.

Due to these reasons, the ambiance needs to be fine-controlled to cope with food deterioration, especially for the packets that have worst shelf life. However, given this complexity, a direct mathematical formulation of optimization problem is impractical, and instead a data driven approach that exploits the central availability of data from throughout the network at the AOC can be adopted. In particular, given the time-series of the data collected from a given carrier (plus data from others carrying same/similar products), one can build models to estimate if the cooling needs to be adjusted to both maintain the quality and satisfy any mandatory constraints.

\subsection{8.4 Gradual and Sparse Infrastructure Deployment}

In the above, we implicitly assumed that every warehouse and carrier in the logistics system operated by a 3PL operator carries our monitoring infrastructure and every box in every pallet has the S&C module. This is unrealistic in practice and does not take advantage of substantial duplication/redundancy in the system. In this section, we discuss how a sparser infrastructure could do an acceptable job. A sparse deployment that considers trade-off between accuracy and cost is also useful for gradual deployment, which is essential in reality.

With sparse deployment, we have certain products that enjoy fine-grain monitoring and data collection by the AOC, while other products are either not monitored at all or monitored in a more traditional way, such as quality assessment and recording of environmental factors (e.g., temperature) only at the source and destination and possibly at distribution centers. Generally, these would be done at a coarse-grain level, such as quality of a small sample of the product and temperature at the pallet level rather than at individual box level. Sparse deployment could also take the form of S&C modules with varying capabilities, e.g., those with simple vs. complex or one vs. multiple sensors and with different communications technologies/ranges.

Given such a heterogeneous environment, a key question is as follows: Given two “similar” items \(x\) and \(y\) each with some measurements of environmental conditions measured at a sequence of strictly increasing pair of observation time instants \((t_{xi}, t_{yi})\), \(i = 1, \ldots, k\) for some \(k\), how similar are they in terms of quality at an arbitrary pair of time instants \((t_x, t_y)\)? Note that \((t_x, t_y)\) could be located in the past or the future relative to the observation points. Clearly, if this problem can be solved, we transfer the quality knowledge from the fully monitored item (say \(x\)) to the unmonitored or partially monitored product (say \(y\)) at any point in time from \(y\)’s journey from the source to the destination.

We assume that the truck/warehouse level monitoring infrastructure (e.g., anchor nodes and local communication hub that connects to the AOC) are provided by the 3PL operator, whereas the inclusion of S&C modules in the boxes is clearly the responsibility of the 3PL clients. In addition, the pallets or containers themselves are expected to have attached environmental sensors (e.g., temperature, humidity, vibration), which are best provided/managed by the 3PL operator. These could well be the same S&C modules that go in the boxes, except that they have environmental, rather than food sensors.

In such an environment, trucks/warehouse rooms with monitoring infrastructure may be offered by the 3PL operator as a premium service to its clients for an extra charge. Furthermore, clients
who do not subscribe to this service may be given an intermediate level service at less cost that exploits the similarity-based analytics mentioned above. The accuracy of such analytics depends on many factors, including the measurement of the environmental factors at the pallet/container level, availability of adequate fine-grain data for the same or very similar product, and the quality check at source, destination, and perhaps at the distribution center(s). Thus, the challenge is to devise flexible analytics that can work with the data that are available regarding the unmonitored product. In the worst case, we only know the following: (a) product type and some indication of condition when harvested, (b) duration and nominal environmental conditions prior to shipment, and (c) expected transit time and nominal environmental conditions for the service that is used for the shipment.

In addressing the similarity problem, the key question is the sensitivity of quality degradation as a function of various parameters that affect quality. As discussed in Section 3, the primary factors are the initial conditions, decay factor (reaction rate), temperature, and time. Unfortunately, none of these are precisely known, and there may be fluctuations in temperature from box-to-box even in a well-controlled environment. Thus, instead of using a direct Arrhenious equation-based estimation, it may be preferable to do a purely data-driven analytics. For each shipment that uses the fine-grain monitoring, the AOC will accumulate time-series data of quality and other information such as temperature, humidity, vibrations, and so on, although the latter may not be on a per-box basis. We can do regression analysis of this time-series data to determine time-based deterioration, and other types of analytics such as clustering or principal component analysis (PCA) to determine which factors are most important for each product type and under what conditions. This can then be used to determine sensitives, and in turn used to estimate quality deterioration of unmonitored shipments based on their similarity to the monitored shipments.

Although one would ideally use the same or very similar product type for this analysis, in some cases there may not be enough data for the product type of interest. In that case, it is possible to ignore the product type and focus on all products that have similar degradation characteristics. This is a more challenging problem and may require choosing products with varying levels of degradation similarity, depending on the sample size requirements. We have developed some similarity search techniques that would be useful in this context. In this work, we assume products with many attributes (e.g., shape, size, weight) and a large list of products where we need to find all products that are similar to a given product on an arbitrary set of attributes and an arbitrary threshold for similarity. Assuming some mechanism to convert each attribute value to a number such that differences in values accurately reflect difference in similarity, the threshold can be defined as an acceptable percentage. These mechanisms along with the notion of “dominating attributes” can be used directly to determine which products to consider for the analytics when there is insufficient data for the precise product of interest.

8.5 Role of Proxy and Indirect Sensing

A true sensing of quality or contamination of fresh foods requires sophisticated chemical or biological sensors used in specific ways depending on the product and how they decompose or react to contamination. However, for the most part, this is unnecessary, and instead most techniques depend on proxy sensing, i.e., sensing of some parameters (e.g., temperature, color, texture) that have only an indirect relationship to the quality or safety. For example, a color change can be triggered by the degradation of chlorophyll or the change of carotene or other factors. The sensing of color can be done using cameras and computer vision technology. The technology is used for green and brown detection for the classification of bananas and tomato ripening. As discussed earlier in Section 5, temperature is easy to measure and has long been used...
as a proxy sensor for quality. However, the temperature within a pallet can vary a lot, as discussed in Section 3, and thus fine-grain deployment is necessary even for temperature sensing.

The change of temperature and concentration of CO$_2$, O$_2$, and ethylene are indicators as well as triggers of quality change. The respiration process makes fresh fruit and vegetables consume oxygen and generate carbon-dioxide, heat, and in some cases, ethylene, which in turn can affect the rate of the reactions. Such correlations can be exploited in choosing which sensors to deploy both from the perspective of quality of sensing and other important aspects such as cost and power consumption. For example, chemical sensors consume much less power compared to the image sensors, and thus the former should provide data more frequently. One can formulate suitable optimization models that include correlations, power consumption, and the sensor cost to derive suitable deployment and sampling mechanisms.

Figure 19 shows a conceptual block diagram of proxy sensing, where the vertices $P_1$, $P_2$, $P_3$ denote the spoilage properties (such as color, smell, gas), and $S_1$, $S_2$, $S_3$, $S_4$ denote different sensors (such as image, gas sensors) that sense the spoilage properties. The edges in between them are the weights that reflect how accurately a sensor senses a spoilage property. For example, an image sensor (or camera) can sense the color change readily but be unable to detect any chemical changes or gas emission. Yet, since the two are related, the color change can be approximately converted to chemical change, thereby resulting in proxy sensing of the chemical changes. Similar arguments hold for contamination sensing. Sensors can be deployed intelligently to cover the desired sensing needs (direct or proxy) with an adequate level of accuracy.

Another form of proxy sensing relates to the similarities between products. Different varieties of the same product or even closely related products (e.g., strawberries and raspberries) will be expected to degrade similarly. Thus, it may be possible to deduce quality of, say, product $Y$ from product $X$. The idea is to do finer grain or fuller monitoring of $X$ than $Y$ without losing much accuracy. For example, if we understand the original quality of $Y$ and pallet-level temperature sensing during its distribution, the finer-grain data about $X$ may be enough to predict quality issues with $Y$. Such a mechanism will not work for contamination, but it may still be possible to use fewer sensors or less frequent sampling of $Y$ as compared to $X$.

We assume that every sensing device has a basic low-power quality sensor, plus other more sophisticated sensors that are deployed sparsely, as described in Section 8.4. The point of sparse deployment is to achieve a balance between cost and accuracy of the quality degradation detection. The key to minimizing the sensors is the ability of the AOC to collect data across the entire logistics network and do correlation and analytics on it. Thus, for example, if we know how a given product deteriorates based on the past data and have some basic sensing (e.g., time/temperature), it may be possible to predict the quality without actual sensors. In particular, we may deploy certain sensors in every $n$th pallet only (for some $n$) and then schedule loading of differently sensored pallets in the trucks intelligently based on the sensors deployed on pallets of same or similar product that are in flight. This can be particularly powerful for proactively controlling contamination and FBI potential at a low cost.
8.6 Security and Privacy Issues

The proposed shared logistics with online sensing, communications, and control turns the T&D supply chain into a de facto cyberphysical system, which brings in a number of security and privacy issues that must be addressed to make the infrastructure acceptable.

8.6.1 Security Issues. Because of the distributed nature of the logistics network and rather slow data rates, the security issues are relatively easy to handle.

Let us start with the local communications between boxes and across anchor nodes and boxes. It is reasonable to assume that the S&C modules are largely trouble-free and uncompromised, as they are managed and deployed by the client in its facility. Also, the compromise of a few S&C modules is not a significant security risk and can be ignored. However, the carriers (e.g., trucks or railcars) are more vulnerable to sabotage, particularly when in unsecured areas. In particular, the anchor nodes can either be replaced or maliciously reconfigured to provide faulty data to AOC. The AOC receives a data stream from each carrier/warehouse, henceforth called an agent, and must run sanity checks and anomaly detection on each stream. The anomaly detection could be based on at least two types of comparisons: (a) comparison of stream characteristics (e.g., statistical properties) of the data stream from the same agent in the past and (b) comparison against streams from other agents that deal with the same or similar products. In either case, it is important to note that it is unlikely that all or most of the anchors will be compromised, which means that the detection mechanism must consider the tradeoff between false positive rate and ability to catch real perturbations. The adversary could take advantage of this and adulterate the data only sparsely. However, because of the difficulty of manually altering the anchors in more than a few agents, the impact on the logistics system is unlikely to be significant.

It is also possible for the adversary to simply disrupt the communications by installing tiny devices in the carrier that emit interfering signals (in case of RF and ultrasound devices) or a strong magnetic field (in case of magnetic communications). However, these are easy to detect at the AOC and unlikely to be significantly disruptive.

The anchor to local hub communications can be easily protected by using low-overhead integrity mechanisms such as the ones we have developed for smart-grid protection applications [88]. The nonlocal communications from agents to the AOC can be easily encrypted, and the AOC can ask for authentication of the agents occasionally to avoid man-in-the-middle attacks [112, 117]. Because of the rather low communication frequency (e.g., once every 15 or 30 minutes), the security overhead is expected to be negligible. Similarly, the communications in the other direction can be secured, and the AOC can be authenticated as well by the agents. A bigger issue is the AOC data breach, which is more of a privacy issue, as discussed next.

It is important to note here that the individual S&C modules and the anchors are likely to develop faults or become inoperational because of the rough environment they operate in. Often it is not possible to distinguish between faulty and compromised modules, and they must be treated identically. In particular, even in the absence of any security threats, it is essential that the AOC does sanity check and anomaly detection [41, 49] on the data received from each agent. The simpler checks may be done in-line, while others are done in the background, perhaps on a periodic basis. Dealing with faulty/damaged S&C modules or anchors also complicates analytics, since the missing/faulty data segments in the stream must be ignored and the corresponding computations done using substitute data from other “similar” streams. Some techniques for sparse data handling become directly relevant in this context.

8.6.2 Privacy Issues. One motivation for why the large retail players have traditionally opted for their own private logistics is that no sensitive information about the logistics operations is known to other parties. Such information could include (a) patronage—from which suppliers the

client gets a given product, and its volume/quality; (b) delivery discrimination—which retailer gets how much at a given quality level, and (c) supply issues when they arise, such as late deliveries, delivery of poor quality or contaminated goods, loss of product, and so on. With the shared logistics, however, it is difficult to keep much of this information from the 3PL operator. In fact, the 3PL operator has the ability to do comparative analytics on the logistics information of its clients and exploit it both for tweaking the services it provides (both in a positive and negative way). Furthermore, it may be possible for individual clients to obtain information about other clients even without any collusion with the 3PL operator.

The situation is quite similar to the cloud computing scenario \cite{69}, where the cloud operator can positively or negatively exploit the usage information of various clients, and the clients may be able to derive some information about the workload that other clients run. In a cloud environment, the only way to avoid revealing data to the cloud operator is to encrypt it using client provided keys. With traditional encryption, the cloud is unable to do any operations on the data, and thus can only provide data storage. However, there has been significant ongoing work on “homomorphic encryption” that allows operations on encrypted data \cite{29, 65}. The so-called “fully homomorphic encryption” (FHE) \cite{71} can allow arbitrary operations, but can be 6–7 orders of magnitude slower than operations in the clear. The “somewhat homomorphic encryption” (SWHE) \cite{42} that allows only certain operations can be done faster but is of limited value in general. As for the metadata, some of it is required by the cloud operator (e.g., size, resource, and latency requirements), some of it would be nice to have for better provisioning/scheduling (e.g., data access characteristics, traffic burstiness), and some of it need not be known to the cloud operator (e.g., the application type or the provenance of the code/data).

In the logistics space, we can think of the quality or environmental monitoring data flowing to the AOC during transit as “data” and the rest as “metadata.” The metadata include basic shipment information (e.g., what product is picked up or dropped off, where, when, how much), product quality/condition on pick-up or delivery, mode of shipment, transit time constraints, environmental requirements (e.g., temperature, humidity), and any other aspects of the service provided by the 3PL operator to the client. One can see in this list all three types of information: required (e.g., product volume, pick-up point), nice to have (e.g., product quality at source and product category), and not strictly needed (e.g., precise product, delivered quality). Note that the information regarding the distribution centers that the product passes through and how long it spends there is for the 3PL operator to decide and may or may not be available to the client.

Given this analogy, if the client does not trust the 3PL operator, the (monitoring) data must be encrypted by the S&C module itself using a client-supplied key (that would not be known to the 3PL operator). Note that, according to our assumption, the anchor nodes and local hub are owned by the 3PL operator and cannot be trusted with client-supplied keys. Because of the battery and computational limitations of S&C modules, the encryption must be very lightweight. This entirely rules out homomorphic encryption, but the traditionally encrypted data at the AOC will be useless, since no analytics is possible on it. An alternative is that the 3PL only provides the anchor and local hub hardware, and the customer leasing the carrier/warehouse room is the one that takes control and configures these devices. In this case, it is possible to do homomorphic encryption at the local hub and correspondingly limited analytics at the AOC. The situation is no different than the cloud doing limited processing in the presence of homomorphic encryption.

Metadata privacy is more easily enforced. In particular, the AOC can carry out much of the analytics without knowing the precise product that is contained in boxes that make up the pallets. Recall that the client provides the pallets to the 3PL operator, and it does not need to reveal the details of the product. The precise sources and destinations can also be masked rather easily,
since the pick-up from individual sources and distribution to final destinations are likely handled by the local logistics that may be outside the purview of the 3PL operator. However, there is a tradeoff—as more information is masked, we need deeper analytics to estimate the quality parameters, and there may still be issues arising from lumping together products with different chemical/biological processes of deterioration. In fact, if the analytics can reliably uncover the masked information (e.g., the precise product and where/when it is procured), this would be considered very undesirable from the client’s perspective.

9 HOW WILL THE PROPOSED ARCHITECTURE BENEFIT THE SUPPLY CHAIN?

The proposed architecture will provide an online infrastructure for monitoring perishability and contamination in the fresh food T&D pipeline, particularly during transport of food on carriers (e.g., trucks) or in warehouses. This involves both the localization of the boxes containing sensor modules that detect quality/contamination issues and its communication to a local hub node. The societal impacts of this research are expected to be substantial: It would help reduce perishable food waste and its carbon footprint while improving efficiencies of distribution and providing food to the needy within the region.

9.1 Efficiency Improvement

The quantitative evaluation of the efficiency improvement of this proposed sensing and communication infrastructure is non-trivial, as it involves deploying the S&C modules, gathering the data, and analyzing them in detail, which is lacking in the current supply chain. However, we summarize some of the literature that has done some preliminary studies and comparisons between the informed and intelligent distribution (which can only be made possible with the proposed infrastructure) as opposed to traditional FIFO-based distribution. The authors in Reference [95] have done a Monte Carlo simulation–based study on cooked ham and have shown that using intelligent distribution policy reduces the amount of loss from 16% to 8% as compared to FIFO-based delivery. Authors in Reference [155] have done a similar study on gilthead sea bream and have shown that the corresponding loss can be reduced from 15% to 5%. A study conducted by University of Florida and Ingersoll Rand [21, 86] found that using quality information-based intelligent distribution reduces strawberry losses from 37% to 23%. Another study from Wageningen University, Netherlands, found that replacing fixed expiration date with a dynamic one reduces the losses of fresh pork from 17% to 4%. These studies are summarized in Figure 20. Also, shipping the packages with uniform shelf life results in easier inventory management and thus lower quality loss and waste.

9.2 Quality Improvement

However, locating food-borne illness is challenging in today’s supply chain mainly due to the following reasons [81]:

1. In the current supply chain, the observations of food contamination occurs when a number of people in a region report illness. Thus, the data regarding individual food items are not available, thus the symptoms of the illness are then linked to individual food items during investigation.

2. The observations available for an investigation are very sparse. Even if contamination travels through multiple nodes in a supply network, it is only observed when some illness is reported in connection to where the contaminated food was purchased (i.e., at the consumption points). The contamination status at the production plant, processing, storage remains hidden in the early stage of investigation.
(3) As the illness is only observed at the consumption points, there is a large distance between the source of the contamination and the observation points. Also, in the current supply chain, the path taken by the products from the origin to consumption points are not always available, which increases the number of all possible candidate paths and makes tracking operations harder.

The most recent case of contamination just happened as of this writing: 32 people from 11 states were sickened due to romaine lettuce being contaminated with E. coli bacteria [27]. The authorities are unable to pinpoint the origin or cause and therefore all romaine lettuce produced anywhere in the country continues to be discarded for almost a week. The latest determination is that it likely originated in California, which would hopefully result in some narrowing of what is discarded.

It is clear that the lack of traceability is very expensive and wasteful. The proposed online sensing and communication infrastructure will greatly assist this process by adopting better tractability as the products move through the supply chain. If the contaminated products are traced even before the consumption points in the food chain, then it will greatly reduce the distance between the source and the detection points and will largely ease locating the contamination origin. The detection process will also be faster because of the abundance of available contamination data and the movement of the packages within the supply chain.

9.3 Deployment Cost Issues
Deploying a comprehensive infrastructure, as proposed here, is a huge undertaking and can only happen organically. There is already a keen realization among all big players in the food industry that technological solutions that go past the current practices can substantially decrease food waste. Current practices essentially monitor and record temperature at the pallet level but generally do not have any communications capabilities. In some cases, the temperatures may be measured at multiple points in the pallet. The quality is usually monitored visually or using cameras only at distribution centers, pick-up points, and end-points when such quality monitoring is possible. Simply recording temperature-quality data for each product type over a long period and doing deep analytics on it can provide important insights into the quality deterioration as a function of...
temperature, initial quality, and other factors such as time of day and weather at the time of harvesting, elapsed time between product pick-up and cooling, and so on. This insight can be used for better product management, including early distribution based on future quality prediction due to temperature variations so far. Walmart has recently started such an effort called Eden [28], which is expected to save them $2B over the next 5 years. Verigo’s Pod quality system [28] is similar but also provides “smart” pallets that not only record but also analyze and then communicate remaining product life to any phone or tablet within ~15 meters. The quality and history of the products is also sent to the Verigo Cloud platform and is visible for further analysis.

The proposed infrastructure also eventually connects to the cloud, and the infrastructure beyond the local sensing and communications is being deployed already. Although the cost of this infrastructure is substantial, it is only a one-time cost and is easily dwarfed by the savings it provides (e.g., Walmart’s $2B/year savings). Even for smaller players, the deployment costs using the commercially available cloud services will be quite modest.

Thus, we believe that the remaining challenges both from technology and cost perspectives are largely in building truly inexpensive sensors and radios that can be embedded in each end consumer-level packages and discarded ultimately by the consumer. This is currently difficult to do, particularly for radios that can reach beyond very short distances. Most of the current sensors are based on MEMS technology, and as discussed in a recent article [63], it is difficult to make devices out of silicon and yet achieve costs of around one cent. The same article then argues that the emerging paper/plastic-based sensors can achieve low costs and at the same time can be much more environmentally friendly. There is a tremendous amount of ongoing research on paper-based sensors [150] and various forms of food quality sensing mechanisms [116]. Technologies such as inkjet-printed on-chip sensors and antennas on flexible substrates, multi-walled carbon nanotube-based solutions are expected to realize thus miniaturization and low-cost solutions. One such example is the C2Sense gas sensor, which costs roughly around 25 cents [159]. Such miniaturized, printed circuit-based solutions are currently being researched heavily [53, 54, 82, 168] and will be expected to help in miniaturization, power consumption, and cost.

While this research finds its way into truly inexpensive implementations, even higher cost sensors can be useful by integrating only one out of \( n \) packages \((n > 10)\). Recall that the proposed mechanism only calls for one sensing and communication module per box and reusability of the sensors. For this type of usage, a cost of less than $1 per sensor and communications module should be acceptable, which we believe is already possible. Even these costs will naturally come down substantially in the next several years, thereby removing cost as an inhibitor.

10 CONCLUSIONS AND FUTURE CHALLENGES

In this article, we discussed how the ongoing rapid transformation in the logistics space can be exploited to provide an intelligent distribution mechanism for fresh food to significantly reduce food waste, enhance freshness, proactively detect food contamination, and improve the efficiency of logistics operations. We provided a comprehensive survey of key issues in this endeavor, including the fresh food deterioration process and its sending, communications technologies for local communications of the sensed data, the use of collected data in improving logistics operations. We also discussed the important issues of gradual and sparse deployment of such a system, and the privacy and security issues brought about by the infusion of information technology into fresh food logistics. We hope that this article will spur new, deeper research into many of the issues articulated in this article.

The success of this architecture largely depends on the rise of low-cost futuristic packaging technologies, along with the development of printed in-packaging sensors and radios to communicate food safety information directly to a centralized repository. Besides developing the sensing
technology, the false positive reactions need to be minimized to avoid unnecessary confusion. However, as the sensing surface needs to be in contact with the food contents, care should be taken to avoid inducing changes to the food from the S&C modules. This can be ensured by using organic bio-compatible materials for the sensors [45]. These technologies along with the rise of tightly coupled industrial Internet of things will make possible the networking, storage, and processing of real-time sensor data across the food supply chain. We believe that with significant ongoing advances in the physical Internet area, automation, sensing, and communications, these approaches are ripe for adoption in the real world.

REFERENCES


Received March 2018; revised April 2019; accepted August 2019