Section 3: Week 7: Radio Frequency Identifiers and Data Mining

Nate Bachmeier

TIM-8130: Data Mining

March 22, 2020

North Central University

# Radio Frequency Identifiers and Data Mining

## Overview and History

Radio Frequency Identifier (RFID) tags provide an economical solution for labeling individual items and tracking their lifecycle. Organizations have been leveraging RFID since World War II, and its popularity has surged due to mass production driving down costs (Bolic et al., 2010). For instance, a retailer can order apples from a supplier then collect per box metrics around the duration to transport and inventory holding times. When a business can monitor the supply chain with precision and granularity, it enables efficiencies across a broad spectrum of scenarios. Consider the costs associated with misplaced inventory as that ties up capital, and once found, those items require discounting. These costs can quickly become factors of magnitude larger than disposable tags, which are only a few cents in volume (Zhang et al., 2018). Alternative solutions, such as bar codes, exist at a lower price point but come with several limitations. A barcode relies on optical readers of static printed patterns, versus RFID does not require line of sight and uses reprogrammable memory. Those enhancements enable advanced scenarios, such as smart containers reporting its unique identifier, the delivery date, and item summary details. A notion of memory segment security exists so that the delivery date cannot be changed—however, removing items updates the counts.

## Project Evolution

Using RFID tags to track inventory through a supply chain, has been mainstream for some time, with many large retailers mandating the practice. More recently, businesses are extending these monitoring pipelines to include floor space and gain insights into their customers' behaviors (Sakuai et al., 2010). Traditional behavior modeling relies on point-of-sale information to understand purchasing decisions. However, with RFID, it becomes possible to track a specific customer picked up a specific garment, tried it on, and did not make the purchase. If the business can understand why the sale did not happen, it can attempt to correct the issue. Perhaps, the desired size was not in stock, and by alerting the customer when more choices are available, it can complete that transaction. Discovering these high probability sales through automated solutions both improves the bottom line and customer satisfaction versus relying on anecdotal evidence. Amazon Go is also changing the status quo by introducing grocery stores without cashiers. Customers can scan into the store, fill up their shopping cart, and *just walk out*. The solution builds upon a combination of RFID, weight sensors, streaming video, and mobile apps to reduce the friction of physical stores, similar to digit checkouts (Wankhde et al., 2018). Polacco and Backes (2018) raise concerns that this technology will negatively impact small businesses and 3.5 million cashiers across the county. However, the mission of an organization is not to protect low-level positions, but instead to provide the best customer service at the lowest cost. Some might argue that reducing the personnel is counterproductive towards that goal. Though tighter integration of solutions like RFID and mobile can deliver consistent expert assistance. For example, a purchaser can simply pick up an item from the shelf and ask the app for reviews, without needing to hunt down serial or model numbers. Buyers have come to expect these conveniences in the digital world, and traditional storefronts need to exceed those expectations. If not, why bother with the hassle of driving to the store?

## Information Generated

Along with metrics around supply chain utilization, insights into incomplete orders, and simplifying checkout experiences—RFID and related technologies can generate information about process inefficiencies and safety issues across the employee population. Norgan et al. (2020) discuss a recent case study where a laboratory transitioned from barcodes to RFID and began automatically scanning all specimens as they moved between locations. Mining the physical path through the builds could discover cross-division dependencies that slowed down processing times. Management could also gain visibility into training gaps, as different work-stations automatically report both the incoming work and employees actualized time. Consider a situation where safely completing a process takes at least sixty seconds, and a technician has a meantime of thirty seconds. Surfacing this information without granular automated metrics would be challenging to even scale across small teams. These learnings apply to other industries and professions, providing employee performance data while reducing leadership overhead. For example, a contracted maintenance crew needs to do one lap around the building every four hours. Instead of paying a supervisor to oversee this effort, RFID scanners distributed around the building can detect the crew is making their rounds.

## Current Limitations

RFID systems can address many entity tracking scenarios, but there are limitations of the underlying technology. Bolic et al. (2010) state that passive ultra-high frequency (UHF) tags use a process called backscatter modulation. This process begins with a scanner emitting a power source that needs to wake up circuitry and generate a query response. That physical process can encounter noise and erroneous values for numerous reasons. For instance, the distance between the scanner and tag can be too far to activate the circuit, or too close and cause nearby tags also to activate. In complex automated environments, multiple scanners are issuing tag queries across the narrow 860-960 MHz band, leading to collisions. When a collision occurs, the protocol needs to detect and reissue the request creating delays that might be unacceptable in specific scenarios. Radio waves run the risk of being detuned as they transmit through fluids and metals, making particular environments incompatible. Similar to other wireless environments, the physical geometry of the room and orientation of tags can significantly impact the performance of transmitting the signal. Because of these issues and related limitations, deploying RFID into an arbitrary scenario is not always possible. Colella et al. (2019) propose leveraging modern technologies like 5G, Bluetooth, WiFi, and Microwave Motion Sensors (MMS). Many IoT platforms offer support for these transports, though at a substantially higher cost per device. The economics of that trade-off depends on the scenario and the frequency of needing to be replaced. Other installations could find that video surveillance provides a more scalable solution. Fong et al. (2016) describe a process using meta-learning to learn gestures and build domain-specific rules around their detection. For instance, a custom gesture could detect that employees are violating safety protocols, and report the infraction. Unlike the previously discussed RFID solutions, the infraction contains video evidence, avoiding any accuracy dispute.

# Conclusions and Final Thoughts

Radio Frequency Identifiers (RFID) provide an economical approach to tag specific objects and track their lifecycle to derive metrics at instance level granularity. Organizations can use this technology to gain insights into a wide variety of scenarios, such as customer behavior analysis, supply chain management, and employee safety. As businesses devise strategies to operationalize this information into business intelligence, it gives them a competitive advantage to reduce overhead and discover partial sales transactions. Amazon Go is evolving these ideas one step further and removing cashiers, allowing for extended store hours and seamless shopping experience. However, RFID has physical characteristics that prevent its deployment into specific environments. Researchers are looking at alternative technologies like 5G and WiFi, to integrate into existing topologies. Irrespective of the underlying media, the process continues to be the same flow with unique identifiers and metadata fixated to each entity. Then decentralized systems can query those tags and make intelligent decisions. When the content is video-centric, alternative representations of those identities can exist, such as through facial and object recognition. As these known agents make decisions within a given environment, rules can predict if their actions are meet some policy. For example, Amazon Go might predict that a buyer took a can of coke from the fridge, and that increases the shopping cart total by 1.50$. These same data mining capabilities should equally apply to the virtual world, such as a web client receives a random identifier and browser cookie. As that client traverse the website, it also creates a path (e.g., click-stream) that analysis can review and optimize. Perhaps the individual in either the physical or virtual world is not particularly relevant. Instead, the aggregate trend becomes the focus of research to reduce specific aspects. Ultimately the virtual, radio, and video scenario are describing the same data mining problem, with the feature space and complexity to train changing.

# References

Adding RFID Capabilities to IoT Technologies: Proof-of-Concept on Microwave Doppler Sensors. (2019). *IEEE International Conference on RFID Technology and Applications. September.*

Bolic et al. (2010). *RFID Systems: Research Trends and Challenges.* Chichester, West Sussex: Wiley.

Fong et al. (2016). Improvised methods for tackling big data stream mining challenges: case study of human activity recognition. *Journal of Supercomputing. October. Volume 72, Issue 10*, 3927-3959.

Norgan et al. (2020). Radio-Frequency Identification Specimen Tracking to Improve Quality in Anatomic Pathology. *Archives of Pathology & Laboratory Medicine. February.*, 189-195.

Polacco, A., & Backes, K. (2018). The Amazon Go Concept: Implications, Applications, and Sustainability. *Journal of Business & Management. Volume 24, Issue 1*, 80-93.

Sakuai et al. (2010). Application of the RFID Data Mining to an Apparel Field. *13th International Conference on Network-Based Information Systems Network-Based Information Systems (NBiS). Sept, 2010.*

Wankhde et al. (2018). Just Walk-Out Technology and its Challenges: A Case of Amazon Go. *2018 International Conference on Inventive Research in Computing Applications (ICIRCA).*

Zhang et al. (2018). Inventory misplacement and demand forecast error in the supply chain: profitable RFID strategies under wholesale and buy-back contracts. *International Journal of Production Research. Aug2018, Vol. 56 Issue 15*, 5188-5205.