The Challenges of Good Data in IoT

ecofit

Peter McGuire | CTO
Ecofit Networks
info@myecofit.com
www.ecofitnetworks.com

ABSTRACT

Achieving long term good data in fitness industry asset management IoT systems requires in-depth knowledge of the assets under watch as well as a continuous monitoring of the IoT systems themselves. Sensors must be shown to be truly tracking the signals they are supposed to track, and any system health failure must be acted on quickly and effectively to reduce lost data. Once at the cloud, correct cloud infrastructure is critical to properly parse, store and act on the mass amount of data that is expected. Ecofit has long term experience in capturing and acting on good data in the fitness asset management IoT industry.

Introduction

Good data should be the goal of any IoT (Internet of Things) solution. Without good data, the real value of an IoT system is impossible to deliver. This seems like an intuitive conclusion: how can anyone draw reasonable conclusions if their data isn't "good" to start with? In reality, gathering good data in an IoT system, particularly a system with a complex implementation in a hostile environment, requires proper system design, continuous monitoring and quick issue resolution. It is our purpose to illustrate this clearly in this paper.

This paper will focus on a particularly challenging industry for IoT solutions: the commercial fitness industry. This is an industry that is very well suited to the value delivered by an automated asset management IoT solution as there are many assets, the assets are high cap-ex, and they are rigorously used on a daily basis. This is a perfect storm/opportunity for high ROI (Return on Investment) IoT solutions. On the flip side, any IoT solution in this industry is strained to deliver good data. An array of issues can arise in the deployment, validation and on-going maintenance of such a system. The industry has a healthy collection of competition. There are many different gym chains, some of them being further divided into franchises, and many individual clubs boast a collection of various manufacturer equipment with distinct needs when it comes to sensors. The "I" part of IoT, the internet, is another challenge with such a diverse set of gyms implementing IT infrastructures with their own rules and philosophies.

On top of these are the broader challenges of any big data system. The cloud infrastructure and database are critical components and must be selected wisely. To deliver on typical IoT value offerings such as real time notifications and long term data analysis reports, the cloud infrastructure must be modern and future proof and the database must be ready for mass time series data input and complex query output.

A Closer Look at Challenges

Mentioned briefly above, there are numerous challenges in collecting, storing and actioning good IoT data in the fitness industry. We will now look a little closer at these challenges and their solutions. We will begin at the sensor and move up the chain to the cloud.

Sensor Fidelity

Sensor fidelity from an IoT asset management perspective is a measure of the ability of a sensor to accurately reflect what is occurring with an asset. In the fitness industry, this defines the ability to indicate whether the fitness equipment (asset) is being utilized or not. Metrics to consider may include utilization, reps, rotations, distance etc. The main focus will be on asset utilization, i.e. accurate usage patterns over time.

Whereas there is an emergence of "integrated sensing" occurring in the fitness industry (where sensing is integral to the equipment and telemetry is delivered via an internet connection) the majority of fitness equipment still requires an externally applied aftermarket sensor to collect asset utilization. These sensors communicate directly with the fitness equipment or capture movement to indicate when the equipment is in use.

The importance of sensor fidelity in these cases is often underestimated, with an assumption that installing the sensor and briefly observing signal tracking is good enough. We have observed a tendency in this early IoT industry adoption to make these assumptions, and this often results in poor sensor fidelity. An initial simple validation is usually not enough, particularly in cases where the utilization metric is the captured movement of the equipment where minimal crosstalk and threshold identification is critical. In cases of direct communication with the fitness equipment using some kind of standard (often the legacy CSAFE standard), the fidelity may be deceptively good at a cursory glance, but can degrade and fail over time.

The solution to these fidelity challenges involve the development of correcting algorithms at the edge that help to ensure good data. Particularly in the case of movement capturing solutions, on the floor identification of predominant movement patterns outside of simple vibration and easy calibration to capture movement patterns is key. When communicating with the equipment directly to capture utilization metrics, an intimate knowledge of communications standards and their edge cases, paired with the ability for installers to fine-tune communications parameters helps to ensure that the communications will persist through usage and power cycles. Regular review of sensor fidelity over time as part of a comprehensive IoT health system also ensures that any failure in fidelity is identified and acted upon to limit data loss. These IoT health systems will be discussed further in this paper.

Connectivity

Good connectivity is the continuous connection of assets to deliver good and actionable data. In the asset management IoT space, a failed connection means a critical failure to deliver the value that was promised by the system. A failure of connectivity could lead to expensive issues like false positives or lack of visibility. Even worse, time and money might be wasted correcting a system that has been already paid for to save time and money.

During our nearly ten years of experience with attaining and maintaining connectivity to ensure good data, we have noted that poor connectivity is the number one cause of bad or insufficient data. For an organization that is not heavily focused on good connectivity, being connected might just mean adding enough functionality to your product to connect to a network. This is not nearly comprehensive enough of an understanding of connectivity to cut it in the delivery of real and lasting value to the end customer. True good connectivity is in fact a long term commitment. It originates at the engineering and design stage, is dependant on focused operations and onboarding, and survives only if closely monitored and maintained. Without these pillars, connectivity will either fail to occur in the first place or will fail prematurely during its lifetime to deliver the value it is meant to deliver.

To help us understand and quantify good connectivity, let's look at two typical examples of IoT solutions in the fitness industry.

Low Latency, Highly Responsive

An IoT asset management solution is sold on the promise of reduced maintenance costs and improved equipment uptime. To realize this value offering, sensors are added to fitness equipment that capture basic utilization patterns over time and deliver this telemetry to a cloud infrastructure. In the cloud, the data is analyzed in real time to detect failure patterns and to deliver a notification when a positive pattern is detected.

In this scenario, good connectivity means ***always on, always available***. If there is a failure in connectivity, the patterns under analysis in the cloud will be skewed and may even result in missing critical events. A good connection in this situation requires continuous monitoring ***on top*** of monitoring the sensor's output. A second layer of notifications needs to be implemented that indicates a failure of the connection itself. This health monitoring is critical to ensure the delivery of good data and will be explored more when we focus on solutions later in this paper.

High Latency, Low Responsiveness

Another scenario is asset utilization tracking. If the value offering of an IoT asset management system is realized through analysis of historical utilization (for example, how much one asset has been used over six months versus another asset), good connectivity looks a little different. To achieve good data, the utilization information needs to be delivered eventually and some latency is acceptable. Good connectivity here can be covered by a sometimes available connection as long as the critical information is recorded and buffered at the edge. A failure might transpire if there is insufficient edge data buffering, or if the latency of data becomes unacceptable. In this scenario it is still good practice to layer on a health monitoring system, however this system need not be as responsive as a situation where low latency is critical. A monthly automated report of bad connectivity might suffice to satisfy the delivery of good data.

From the examples above, we can draw the following metrics of good connectivity: ***Acceptable latency, Edge buffering, Health of network***

As acceptable latency goes up, the need for a very healthy network goes down and the requirement for good edge buffering goes up. If acceptable latency is low, the opposite is true. Both scenarios require ongoing connectivity monitoring and attention.

Connectivity Health Monitoring 

An immediate solution to failed connectivity is an automated health monitoring system that will constantly monitor whether local wireless networks and gateways are running and, if possible, assign a "connectivity score" to both connected assets, the wireless network, and any gateways in the system. This information can be used for real time notifications of network failure as well as a confidence margin on data captured to use when generating reports or real time customer facing notifications. If the score is low, any report or customer facing notification needs extra scrutiny.

Deployment and Maintenance Operations

The logistics of deploying and maintaining IoT systems that deliver good data is heavily dependent on strong logistics and operations procedures. This is particularly true in the fitness industry, with its unique challenges to obtaining and maintaining good sensor fidelity and connectivity. As mentioned in the beginning of this paper, assets are in a harsh environment and behind some variation of firewalls and IT rules.

We have talked about the importance of engineering already, with deep knowledge of edge cases, well designed installation apps that limit expectations of installer crews, and continuous monitoring systems being key concerns. These solutions must work very closely with operations teams in a feedback cycle where deployment and maintenance concerns are brought back to engineering and considered for perpetual good data improvements. This development cycle is on top of the development delivering the value offerings themselves, but is critical to the delivery of good data.

The operations team is very important. When a notification fires from the health monitoring systems, it is key that some action is taken.

Our experience has shown that a solid customer onboarding process is needed to not only bring the customer up to speed with their product, but also to identify proper customer departments that can be contacted in the case that the IoT system is failing to collect good data (outside of the IoT system's own potential failures). The role of operations varies from customer to customer, but the objective is typically the same: Engaging the customer and identifying departments or people within the customer's organization who will be able to act on the ground to correct blocking items such as IT firewalls, sensor misplacement, and gateway damage or disconnection. Without these operations relations in place, a drop in connectivity or sensor fidelity can cause serious damage to good data.

Another key requirement of good data is good contextual data to support it. In the fitness industry, it is critical to know all relevant details about the equipment:make, model, add-ons and so forth. Operations that pay attention to these details will enrich the contextual data to accurately reflect the equipment on the floor and help to keep the results of the data analysis relevant.

Scalable and Dependable Cloud Infrastructure

When there is good deployment, sensor fidelity and connectivity from the asset all the way to the cloud there are important best practices that need to be followed in order to handle the massive influx of data, process the data in a timely manner, and store the data in the right style database for the job. If logic and storage components in the cloud fail under massive scale or downtime, this can mean lost data.

Use of more modern managed cloud solutions helps to shift the burden of scalability and uptime to the service provider, who are more than ready to take on this challenge (AWS for example). Managed and serverless solutions not only ensure massive scalability and uptime, they also tend to keep prices down.

A longer term concern for IoT time series data collection and storage is the choice of database. IoT is still a young industry, and traditional data storage has had to catch up with its unique requirements. A proper database for the job needs to be able to take in a large volume of data over a short time and store it securely and dependably. Also, the database needs to perform well under the complex queries that are necessary for drawing useful conclusions from the data. A new breed of databases focused on time series data has emerged, which tick all of the boxes needed. An added bonus is that many of these solutions are built on top of traditional SQL style databases, meaning the development learning curve is low.

Solutions In Depth

Solutions to the challenges of good data collection have been alluded to throughout this paper, but let's take a closer look at them so they are easily understood.

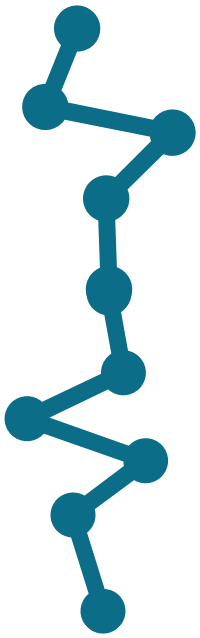
Sensor Fidelity

The solution we at Ecofit have devised to ensure good sensor fidelity is multi-faceted. First, a strong understanding of the protocol limitations between sensor and equipment is built into how the sensor communicates with the equipment. CSAFE is the primary communication protocol supported by fitness equipment in the field, and this protocol is very unevenly implemented across brand and model. To overcome this, Ecofit uses a very limited implementation of the protocol that covers the greatest population of CSAFE enabled equipment. There are some limited configuration differences needed for installation, and the process to identify which configuration is needed is very user friendly and straightforward.

In the case of non-CSAFE equipment, a motion sensor is applied. The pressing concern with motion sensors is that they pick up the usage of the equipment properly. This is a balancing act: too sensitive a threshold and the sensor will pick up vibrations not related to equipment usage. Not sensitive enough and slight usage will not be picked up. A solution to this is two fold. One, identify movement patterns such as movement along an axis and circular movement. Two, make the installation calibration very simple. If the calibration process is complex, it ties up the installation process and the chance of poor calibration goes up.

Connectivity

As noted, a key component of long term connectivity is the continuous monitoring of the health of the connection. An automated system that will score the connectivity of an installation from the local network up to the cloud should be in place. This system should notify users when connectivity is poor so that attention can be paid to the problem and the connectivity can be brought back up to appropriate levels. In the case of connectivity, some solutions require high availability and some do not.



It's up to the designer to decide what is appropriate connectivity given business objectives.

Deployment and Maintenance Operations

Good deployment and maintenance operations is key to any good data delivery system, especially in the fitness industry. Growing and maintaining good contacts with the right people in each customer organization means when issues arise they can be solved in a timely manner.

Scalable and Dependable Cloud Infrastructure

As mentioned, the right cloud infrastructure and storage must be implemented. A cloud infrastructure that promises to scale and is highly available is critical to maintain the data paths running. The choice of storage should be optimized for IoT time series data, be highly available, scale appropriately, and properly serve the data in various high load queries.

Conclusion

Good data is of course critical to the extraction of value that IoT systems are meant to deliver. As indicated in this paper, the collection of good data in the fitness industry is more challenging than it may first appear. Close attention needs to be paid to the health of the IoT system, network connectivity, and the receiving cloud infrastructure. Well designed systems also should integrate closely with operations teams to assure the immediate resolution of issues affecting the collection of good data. Our experience at Ecofit has taught us these important lessons and has guided our design and organizational structure. If you are interested in benefiting from this experience, please contact us at info@myecofit.com and request a demonstration.

ecofit™

ecofitnetworks.com | info@myecofit.com