Lost in Space: RFID Object Localization in a Dynamic and Noisy Environment

ALEXANDER LAMSON, University of Massachusetts Amherst
EMMANUEL CECCHET, University of Massachusetts Amherst
PRASHANT SHENOY, University of Massachusetts Amherst

An astronaut's time is very expensive, and they have a lot of cargo to keep track of. To ease their burden, cargo can be localized remotely using Radio-frequency identification (RFID). To accomplish this, RFID tags are placed on Cargo Transfer Bags (CTBs). Fixed antennas and marker tags are mounted on the walls at known locations and are used to calibrate the localization algorithms.

The unique environment of the International Space Station (ISS) poses two serious challenges to RFID localization. First, it is dynamic. New tags are added to the station, tags move around within the station (sometimes to uninstrumented areas) and antenna readers get turned off both intentionally and unintentionally. Second, it is noisy. Many RFID reads have invalid Electronic Product Codes (EPCs), and there are many metal objects causing multipath reflections.

This paper explores three years of historical RFID data, from 2017 through 2020. Using this data, we create an overview of the history of the tags and characterize the challenges in localizing the tags.

CCS Concepts: • **Information systems → Sensor networks**.

Additional Key Words and Phrases: Applied databases, Internet of Things, RFID

1 INTRODUCTION

NASA uses the RFID Enabled Autonomous Logistics Management (REALM) project [4] to automatically track and localize inventory with RFID tags and antennas. Object locations are represented using XYZ Euclidean space.

REALM is divided into REALM-1, REALM-2 and REALM-3. REALM-1 uses fixed antennas to read RFID tags attached to objects in the space station. REALM-2 uses a mobile drone named "Astrobee" to fly through the ISS and take RFID "snapshots" to refine item localization. REALM-3 uses "smart drawers" where antennas are fitted to drawers to further increase localization accuracy. This paper uses the data collected during the deployment of the REALM-1 system.

The IV&V Management System (IMS) takes a different approach from REALM, relying on manually entered information in order to monitor inventory. Astronauts manually scan barcodes of items they move around, and log them into IMS. IMS uses a tree structure to represent where objects are stored (i.e. a tagged CTB is stored in a rack, which is found in a module). This process requires manual human involvement, which results in data that is sparse, tedious to maintain and error prone. It is estimated that the information logged in IMS is approximately 80% accurate. For the purposes of this paper, only objects tracked using the REALM system will be considered.

Although some human-related challenges are solved through REALM's approach, automatic localization itself remains a difficult problem, made more challenging by the properties of the ISS:

- The system is dynamic:
  - New tags arrive onboard visiting vehicles
  - Old tags are burnt in the atmosphere or returned to Earth on a visiting vehicle
  - Tags move within the station
  - Antenna readers turn off both routinely and unexpectedly
- The readings are noisy:
  - Some readings have invalid EPC numbers

- Antennas are reading tags inside of a large metal tube which causes multipath reflections
- Objects and astronauts may sometimes obstruct tags and antennas

Finally, not every module in the ISS is fully instrumented, as can be seen in Figure 1. This creates the potential problematic situation that a tag moves from an instrumented area to an uninstrumented area during a reader outage. When this occurs, it is impossible for the system to localize the lost tag, and its location is assumed to be where it was last located.

2 BACKGROUND

RFID has been used in various commercial applications. Examples include supply-chain management (ex. automatically tracking warehouse goods), access control (ex. key cards to enter restricted buildings), and payment (ex. E-ZPass for highway toll collection). [6]

An RFID system consists of a few readers with antennas attached to them, as well as tags attached to the objects to be localized. The readers, which can be either stationary or mobile, communicate with the tags within its wireless range and collect information about the tags. In this case, the information collected from a tag includes the time the tag was read, the tag's identifier and received signal strength. Tags are classified into three categories: passive, semi-passive, and active. Passive tags, which are the type used on the ISS, have no internal power source. They use the electromagnetic field transmitted by a reader to power their internal circuit. In order to transmit data back to the reader, tags use backscattering rather than a transmitter. Semi-passive and active tags require a power source and are not used in the REALM system. [6]

A particular property of RFID is multipath, where the RF signal can be received via more than one path. This occurs because RFID uses waves which reflect off of materials like metal and water. This results in two effects. First, in some cases the reflections allow the antenna to weakly read tags which it doesn't have line-of-sight to. Second, received reflected signals can interfere with each other, resulting in a weaker received signal strength. This form of signal attenuation is called multipath fading. A weak signal strength cannot always be attributed to this phenomenon, though, because signal attenuation can be caused due to distance, as the signal weakens farther from the reader.

Before the REALM project, astronauts could only track objects on the ISS by scanning barcodes on the objects as they use them or put them away. This requires precious astronaut time. The main benefit of RFID is that it does not require human interaction, and would therefore save astronaut time which should be used as efficiently as possible.

3 SYSTEM AND ENVIRONMENT OVERVIEW

The total internal pressurized volume of the ISS is about the same as a Boeing 747, at approximately 1,000 cubic meters. [3] 118 cubic meters of that space is usable stowage space. [9] The ISS is made up of 16 pressurized modules (also called "nodes") [2]. All antenna instrumentation is located in the 3 sequential central modules NODE1, USLAB and NODE2. This instrumented area is 20.6 meters long. The 6 adjacent modules JPM1, COL1, NOD3, A_L1, PMA1, PMM1 have limited coverage. In other modules located on the periphery of the station, coverage is too limited to perform localization. This coverage is visualized in the map of the ISS in Figure 1.

There are 6 readers, with 4 antennas connected to the ports of each reader, for a total of 24 antennas. [1, p. 35] Each reader acts as a hub, taking in the information from the antennas connected into it. Readers are custom designed by NASA, and are based on the Jadak ThingMagic M6e reader module. [9] All antennas use passive ultra high frequency (UHF) RFID, in the 900 MHz frequency range. Two types of passive tags are used. The first is Alien Technology's

Module	**Marker Tags**
NOD2	33
NOD1	32
LAB1	29
PMA1	13
NOD3	6
Total	**118**

Table 1. Total number of marker tags per module. Unlisted modules contain no marker tags.

Squiggle, which is used on non-conductive items. The second is Metalcraft's Universal Mini, which can be placed on conductive items. All marker tags are the Metalcraft type. [9]

Antennas are located in the hatches between the modules and are oriented to provide coverage over the modules. For example, on reader 3 port 2, antenna 12 is located in the hatch of USLAB (connecting to Node 1) and is pointed in the direction of Node 2. This means this antenna can read tags down the length of USLAB, some of the tags located in Node 2, and it can also weakly read some of the tags located near and behind itself due to multipath reflections. The visibility of different modules per antenna can be seen in Figure 3.

Objects to be tracked are often placed in white fabric duffle bags called CTBs. A CTB will typically be assigned a single RFID tag. CTBs can be found in racks, where they are held in place with elastic cords. Tagged objects can also be found in drawers.

In addition to the RFID tags placed on the CTBs that we are trying to locate, there are RFID marker tags placed on the walls of the ISS. These marker tags are placed at known locations, which provides ground-truth information about the signal strength to each antenna for these locations. The number of marker tags in each module can be found in Table 1. The majority of marker tags are located in the three instrumented modules, with an additional 19 marker tags located in PMA1 and NODE3.

Rockets are used to send astronauts and supplies up to the ISS and back to Earth. A ship that docks with the ISS is called a "visiting vehicle". A new ship docks on average every 48 days, and stays docked on average for 51 days. [7] Rather than unloading all the CTBs from the docked ships immediately, items are taken from the docked ship as needed over the period of about a month. This means that there is typically not a large spike in the number of tags first seen when a ship docks.

The purpose of measuring the quality of ground truth and the specifics of the system onboard the ISS is to support and optimize the performance of the localization algorithms actually used to find the RFID tags. The REALM system uses three distinct types of localization algorithms.

1. Algorithms based on antenna location and orientation, as well as signal strength
2. Algorithms based on marker tag signals
3. Machine learning algorithms trying to combine all the info available

All algorithms rely on ground truth (known antenna and marker tags locations), but there are still sometimes mistakes. Although IMS could potentially be used as a source of ground truth, its accuracy is also too low to be fully trusted.

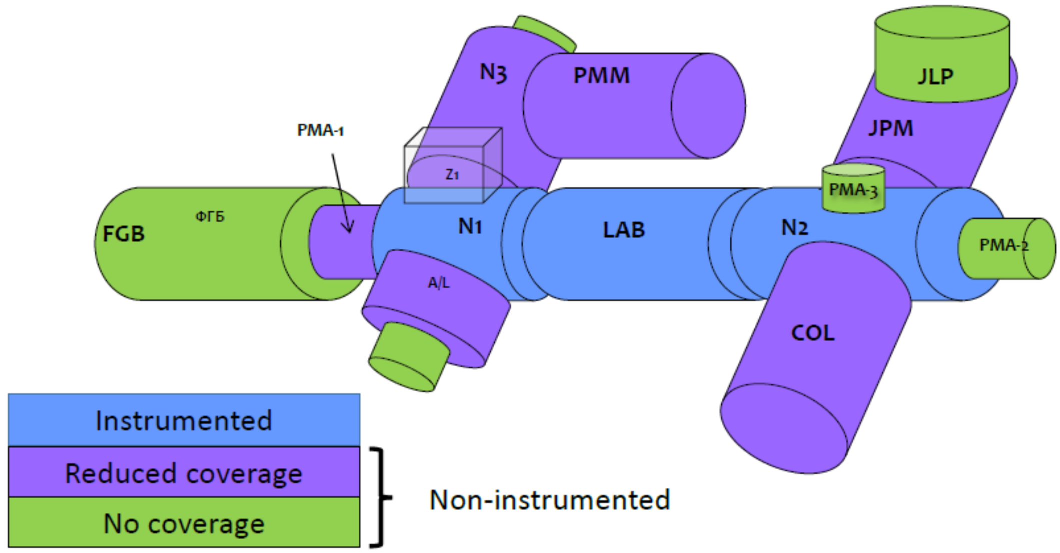


Fig. 1. Map of the space station. Blue regions denote modules instrumented with antennas.

4 DATA ANALYSIS

4.1 Dataset

The dataset comes in the form of raw RFID readings in a database. The data collected spans from 2017-02-14 through 2020-03-31. Each row of data collected by the antennas contains the following columns.

$date$ Timestamp when the tag was read, with millisecond granularity.

epc_id Unique EPC identifier for the RFID tag. There are a total of 11,748 distinct valid EPCs.

$antenna\_id$ Identifier number for the antenna which read the tag.

$rssi$ Received signal strength indicator of the tag, measured in decibels. Values fall in the range of -90 (weak) to -30 (strong).

$freq$ Frequency of the signal from the tag. Values range from 902.750 to 927.250 MHz.

$phase$ Phase of the signal from the tag in degrees.

$power$ Power of the signal from the tag.

cnt Number for reads from the tag. Multiple read hits are aggregated under this field.

It's important to note that RSSI (Received Signal Strength Indicator) does not correlate perfectly with the distance between the tag and the antenna. This is because there is more than one reason why a tag's signal strength could be weak. The signal could be coming from a direct line-of-sight hit of the tag that is simply attenuated due to distance, or it could be that the RF wave's path to the tag was reflected, in which case the tag could actually be physically close to the antenna. It's also possible that the tag or the antenna was partially obstructed by a CTB or other obstacle, which would result in an all-around drop in signal strength.

Additionally, it cannot be assumed that frequency and phase remain constant, even for tags which are not handled by astronauts. This is because under microgravity, tagged objects that aren't fixed in place can freely rotate and drift.

4.2 Best-case RF Environment Characterization

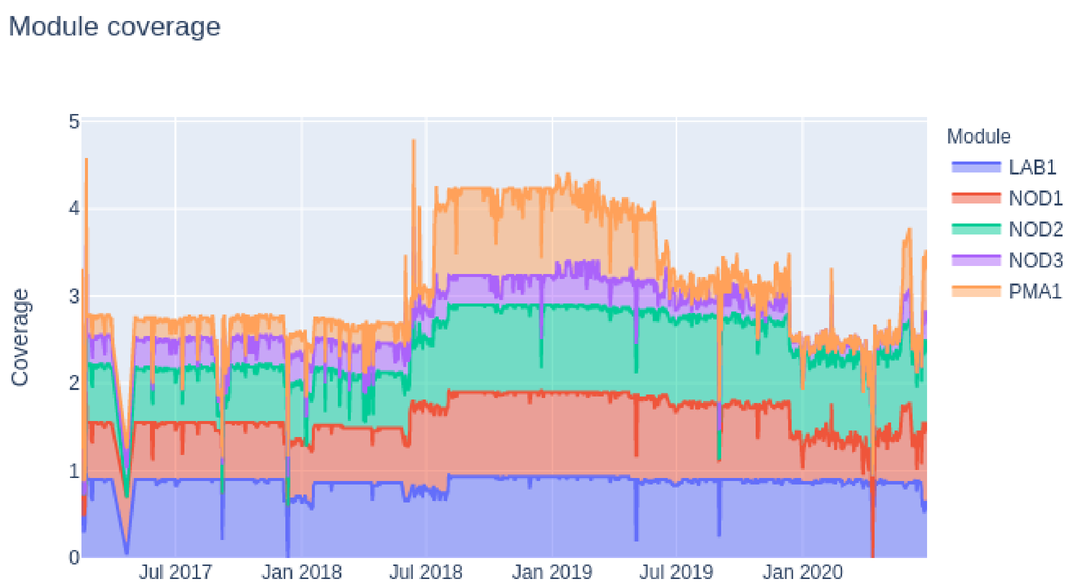


Fig. 2. This plot visualizes the percentage of marker tags visible in each of the instrumented areas. One unit on the y-axis represents 100% visibility of markers tags in that module.

As is shown in Figure 5, readers can turn off unexpectedly for long periods of time. Additionally, readers can turn off for shorter periods of time during the course of a day. It would be unfair to include this unplanned downtime when characterizing the nominal or best-case performance of the system. Therefore, a timeslice must be taken which most closely approximates the system operating with little to no unplanned downtime. In order to find this timeslice, we plot the module coverage over time in Figure 2. The coverage is measured as the percentage of marker tags visible in that module at a given time. This gives an approximation of how much of the ISS is visible at any time over the three years of data. We find that the system can see the most during the weeks from 2018-08-20 though 2018-09-16 (approximately a month). This period is visible at the tall, flat region in the middle of Figure 2.

If we zoom in on that period of nominal weeks, we can visualize how much of each module each antenna sees in the nominal case. This visibility is illustrated in Figure 3. We find that each antenna can see a subset of the modules, but no one antenna can see all the modules. For example, the antennas on reader 5 (antenna IDs 27-30) have a very high collective visual on the marker tags in NODE2. This means that we can have an idea of which modules will have reduced coverage in the event of a particular reader going down.

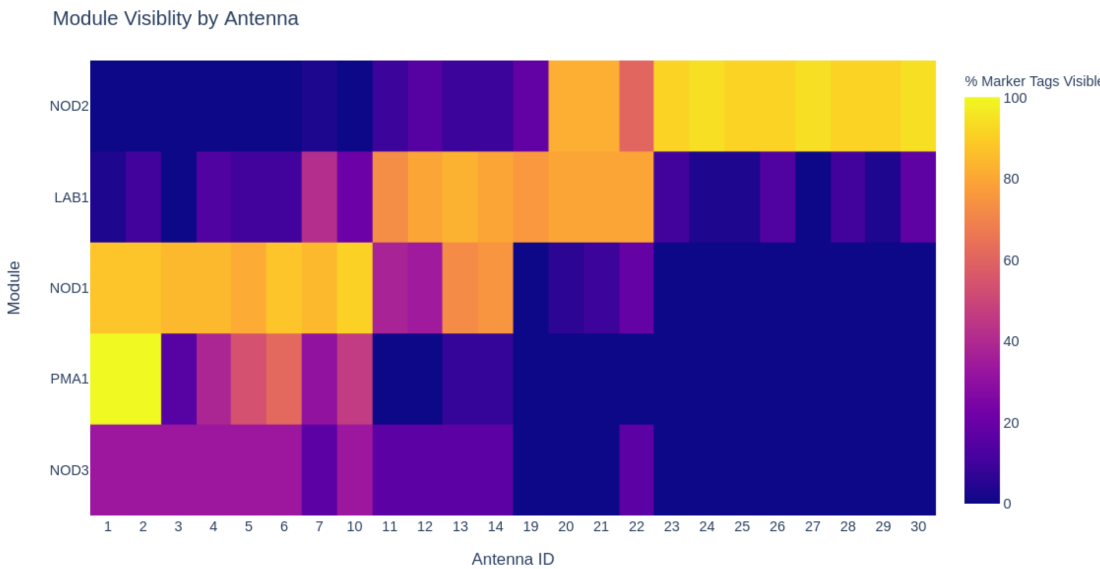


Fig. 3. Visibility of the instrumented modules by the antenna id.

Looking at these nominal weeks, we typically read between 1,000 to 1,500 distinct tags, or an average of 1275.2 distinct tags per day. We nominally receive on average 931,931 reads per hour, or about 21 million reads per day. Of those tags, an average of 101.9 of them are marker tags. This accounts for 86.3% of the total 118 marker tags.

A small percentage of these daily reads have invalid EPC numbers. More depth is given in Section 4.2. These bogus reads are found to be uniformly distributed over time, and continue to occur on an identical reader setup on Earth.

A daily average of 181.1 tags are seen for the first time, and 180.9 are seen for the last time.

The system achieves on average 85.28% uptime per day, which is 20.5 hours of uptime and 3.5 hours of downtime. Here we define the uptime percentage as the total minutes with reads across all antennas, divided by the maximum possible total minutes with reads across all antennas.

Even in the situation of a fully operational system, there are challenges. The system must be turned off multiple times per week while the ISS undergoes safety checks. These scheduled downtimes last for 4 hours each. During this downtime, all antennas are turned off, resulting in a loss of historical data. Additionally, even in the case of operational antennas, there are still unplanned outages. On the average nominal day, there are 70 minutes of planned downtime and 141 minutes of unplanned downtime. Finally, there are constant bogus reads which must be filtered out.

5 PLANNED AND UNPLANNED DOWNTIME

The system is not 100% operational at all times. The system can be in one of three states: fully up, partially up, or fully down. In the fully up state, all antennas are reading tags properly. In the partially up state, a subset of readers fail to read any RFID tags. In the fully down state, no antennas read any tags.

Because RFID tags are most commonly attached to objects which are deeply stored away for later use, antennas will often not be able to read them until the object is removed from storage. This means that if a downtime occurs while an object is removed from deep storage and re-stored in a location where the antennas cannot read the tag, then localization will be impossible. Additionally, if a tag can be seen as the system goes down, but is placed in deep storage during the downtime, then the system will only know that the tag was moved, and not where it was moved to.

Over the past 3 years, 25.61% of days the system has been 100% operational (defined as >50% average uptime over all antennas per hour of the day). The data shows that entire readers turn on and off, but individual antennas do not. We want to maximize uptime and minimize planned downtime, which we can improve on if we know how much there is over time. To visualize this information, the weekly downtime was categorized for each week in Figure 4. We find that as the planned downtime schedule changes over time, the total amount of planned downtime decreases. However, due to reader outages, the total amount of uptime has also decreased. This means reader outages and unplanned downtime are currently the biggest causes of downtime.

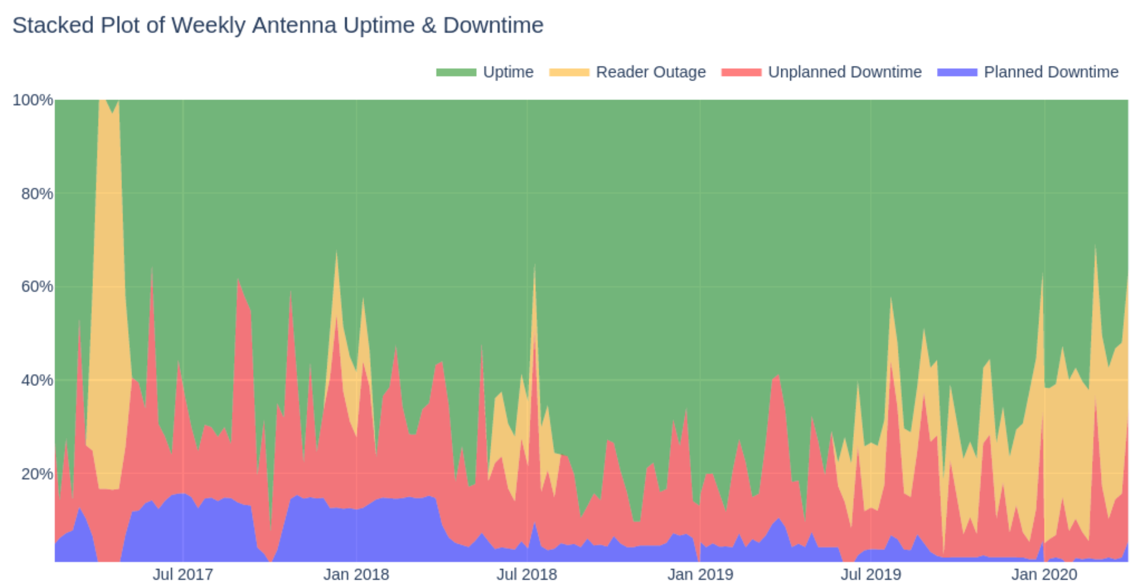


Fig. 4. Stacked plot of how time is spent each week.

5.1 Unplanned downtime

Unplanned downtime can occur for various reasons. CTBs, astronauts, water and tools can all obstruct antenna readings. If an obstruction is in front of a tag, then only that tag will be affected. However, if the obstruction is in front of an antenna, the number of reads from that antenna will be reduced or stop completely.

Reader outages can also occur due to accidents. For example, reader 1 experienced a power issue, causing it to repeatedly reboot unexpectedly. Due to this problem, on 2019-05-31 it was removed from power until it could be replaced or repaired. This configuration is currently ongoing, and is an example of long unplanned downtime. We see this power issue visualized in Figure 5 as a reader outage. As is shown in this figure, every reader has had an outage at some point in time. The speed at which a reader can be brought back from an outage depends on the difficulty of the repairs, as well as availability of the astronauts to perform the repair - this is why some outages last longer than others.

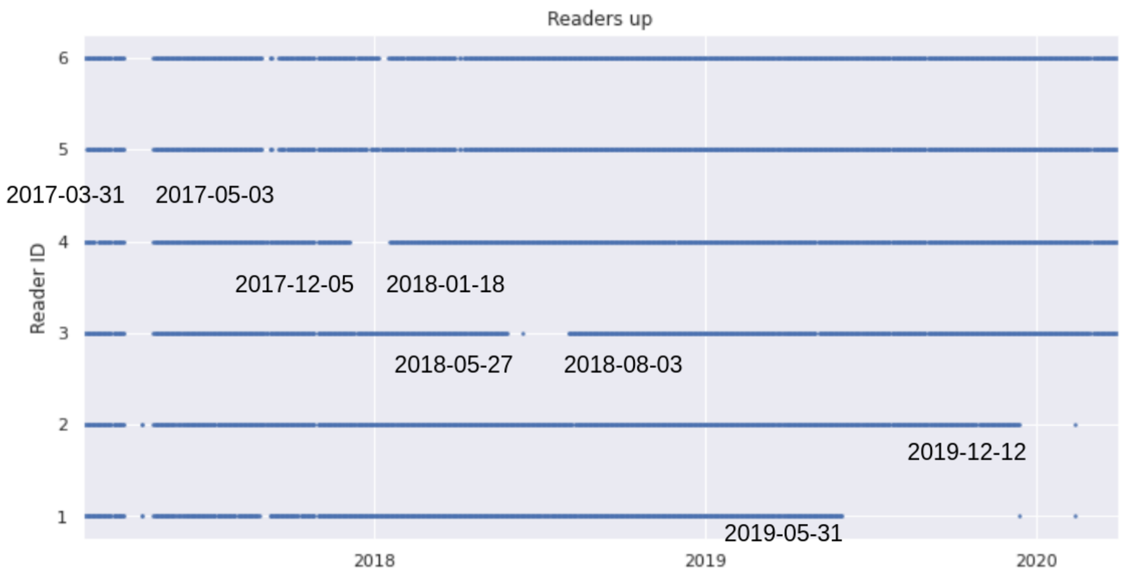


Fig. 5. Strip plot of the uptime of the different readers. A dot is placed when the reader is online at least half the day.

5.2 Planned downtime

The structural integrity of the ISS is checked regularly to ensure the safety of the astronauts. During these checks, all antennas must be shut down to prevent interference. These are planned periods of antenna downtime. The procedure for these planned downtimes has changed over time, and there are three distinct phases.

Planned downtime is special because it can be controlled. Having planned downtime occur less frequently results in more complete historical data which improves localization. Additionally, the time of day that planned downtime occurs can be controlled to minimize the amount of unrecorded tag movement. This will increase the probability that a lost tag can be found.

Below are the phases of planned downtime schedules. These schedules are also visualized in Figure 6.

Phase: 1 **2017-02-15 to 2018-04-02** Readers turned off every day from 3am-6am

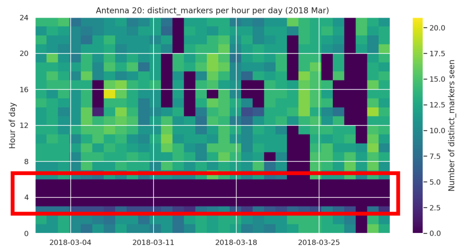
Phase: 2 **2018-04-03 to 2019-09-03** Readers turned off 2-3 times per week from 3am-6am

Phase: 3 **2019-09-04 to 2020-03-31** Readers turned off every Tuesday from 1pm-5pm

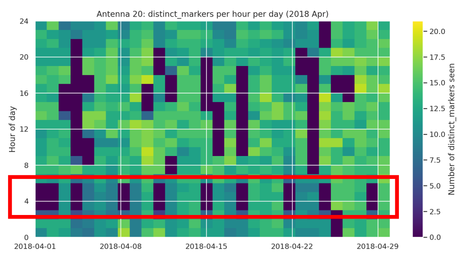
6 BOGUS READS

Each RFID tag has a uniquely identifying EPC. The EPC is formatted such that it is exactly 24 hexadecimal characters long (12 bytes), and always ends in four zeros. In this way, if an EPC has more or less than 24 characters, or doesn't end in 4 zeros, it is an invalid EPC. Many tag reads are actually bogus, as can be determined by checking the EPC of the reading for validity. On an average nominal day, 177 reads are bogus, which makes up 0.00088% of reads.

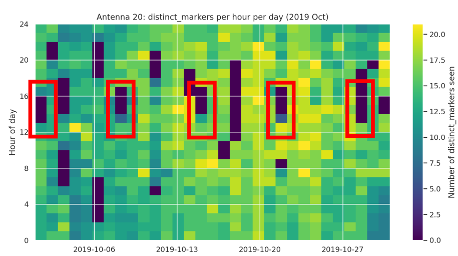
The space of possible invalid EPC codes is large, so 98.9% of the time invalid EPCs are only read once. This unique characteristic can be used to quickly approximate the amount of bogus reads during a window of time.



(a) Phase one planned downtime:
Readers turned off every day from 3am-6am



(b) Phase two planned downtime:
Readers turned off 2-3 times per week from 3am-6am



(c) Phase three planned downtime:
Readers turned off every Tuesday from 1pm-5pm

Fig. 6. The three phases of planned downtime. The x axis represents each day, the y axis represents each hour of the day, and the color intensity represents the number of marker tags read in that hour. Each pattern of planned downtime has been highlighted with red rectangles.

7 ANOMALIES IN GROUND TRUTH

Localization algorithms rely on various forms of ground truth: antenna location and orientation, marker tag location and/or IMS localization information. Detecting errors in these ground truth sources is very important as they can negatively affect the accuracy of the localization algorithms.

Despite best efforts, mistakes can be made. Past mistakes include:

- Plugging antennas into incorrect reader ports
- Relocating an antenna
- Pulling on a reader's power cable resulting in a reader outage (loss of four antennas)

These sorts of problems can negatively affect the accuracy of the localization algorithms. Therefore, detecting these issues in the ground truth assumptions can improve the overall performance of the system despite imperfect conditions.

One method to find potential anomalies is to plot the number of marker tags seen per hour for each antenna, and look for any sudden peaks or valleys, which would mean that could be an anomalous event at that time. We plot the number of marker tags seen per hour for antenna 19 in December 2018 in Figure 7. In this plot we can see a large spike in the number of distinct tags read per hour on antenna 19 on 2018-12-27. After checking what happened physically at that time, it was reveals that the spike was caused by antenna 19 being moved upwards by less than a meter. This increase in reads occurred because a large sleeping bag was no longer obstructing the view of the antenna.

Although an increase in reads generally means improved localization accuracy, the movement of an antenna requires a recalibration for the algorithms, and in particular algorithms which do not use marker tag data for calibration. Future work could explore automatic detection methods for these changes in ground truth.

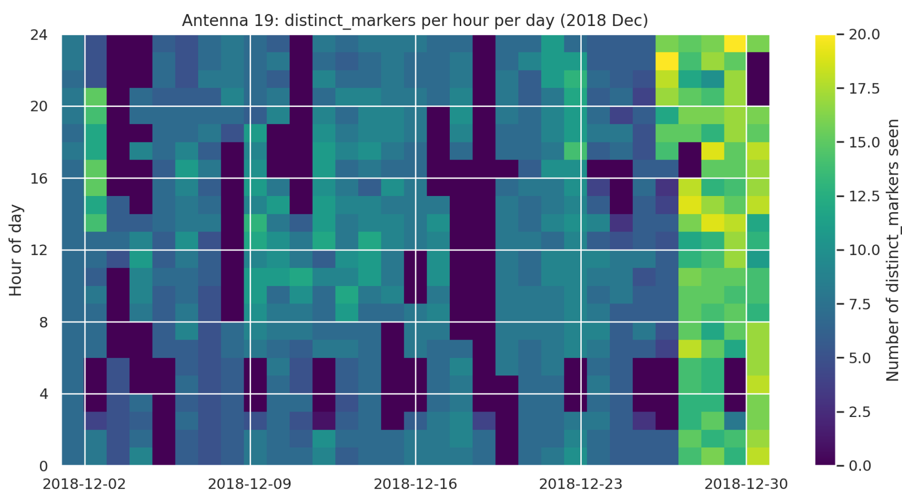


Fig. 7. Large spike in the number of distinct tags read per hour on antenna 19. The spike is visible on 2018-12-27 at 17:00.

8 RELATED WORK

Jue Wang et al. [10] present a work which was the first fine-grained RFID positioning system that was robust to multipath and non-line-of-sight scenarios. They did this by exploiting the local similarity of the multipath environment between nearby tags to create "multipath profiles". However, this technique relied on antenna motion in order to capture these profiles, which is not possible in the ISS REALM-1 system.

In 2013, Chawla et al. [5] showed that RSSI can indeed serve as a reliable metric for localization.

Jing Wang et al. [11] also uses marker tags, but it uses them to establish a relationship between inter-tag distance and phase information. The localization technique used here shares some similarity to what is used in the REALM project.

In the work of Schneegans et al. [8], RFID is used for robot self-localization in an indoor environment. This paper used a particle filter for rapid localization, rather than a slower approach that relies on historical sensor data.

9 CONCLUSION

This paper provides an overview of the data collected by the REALM project used by NASA to localize missing objects in the International Space Station. This project is important because automatic localization will enable astronauts to save time – which is very expensive – that they would otherwise spend looking for lost items. This problem consists of many challenges, several of which are unique to the ISS. These challenges include unscheduled reader downtime, scheduled reader downtime, tags moving to uninstrumented areas, reads with invalid EPC identifiers, multipath reflections and signal attenuation due to obstructions. These properties result in the RFID data that is incomplete and messy. Despite these challenges, much can be learned by analyzing this data.

An analysis of the system's downtime led to the classification of these periods into planned downtime, unplanned downtime, and reader outages. Planned downtime was further characterized by three distinct phases of downtime schedules over the three-year period. Additionally, bogus reads with invalid EPCs occur daily and must also be carefully identified and filtered from the data. Finally, this system exists in the real world where human errors and anomalies occur, so such events must be discovered and accounted for. Revealing these properties of the data from RFID localization in the ISS illustrates the unique challenges of this problem, which can help inform those who do future work in finding objects lost in space.

REFERENCES

[1] 2017. Cislunar Habitation & Environmental Control & Life Support Systems. https://web.archive.org/web/20200801180408/https://www.nasa.gov/sites/default/files/atoms/files/20170329-nacheoc-crusan-gatens-hab-eclss-v5b.pdf. Accessed: 2020-08-19.

[2] 2019. Space Station Assembly. https://www.nasa.gov/mission_pages/station/structure/elements/space-station-assembly

[3] 2020. International Space Station Facts and Figures. https://www.nasa.gov/feature/facts-and-figures. Accessed: 2020-08-19.

[4] 2020. NASA TechPort. https://techport.nasa.gov/view/93175

[5] Kirti Chawla, Christopher McFarland, Gabriel Robins, and Connor Shope. 2013. Real-time RFID localization using RSS. In *2013 International Conference on Localization and GNSS (ICL-GNSS)*. IEEE, 1–6.

[6] Vipul Chawla and Dong Sam Ha. 2007. An overview of passive RFID. *IEEE Communications Magazine* 45, 9 (2007), 11–17.

[7] Mark Garcia. 2015. Visiting Vehicle Launches, Arrivals and Departures. https://www.nasa.gov/feature/visiting-vehicle-launches-arrivals-and-departures

[8] Sebastian Schneegans, Philipp Vorst, and Andreas Zell. 2007. Using RFID Snapshots for Mobile Robot Self-Localization.. In *EMCR*.

[9] Claire Swedberg. 2018. RFID Aids Space Exploration on Mars and Beyond. https://web.archive.org/web/20200827171522/https://www.studio98test.com/ahmad-test/rfid-aids-space-exploration-on-mars-and-beyond

[10] Jue Wang and Dina Katabi. 2013. Dude, where's my card? RFID positioning that works with multipath and non-line of sight. In *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM*. 51–62.

[11] Jing Wang, Yongtao Ma, Yu Zhao, and Kaihua Liu. 2015. A multipath mitigation localization algorithm based on MDS for passive UHF RFID. *IEEE Communications Letters* 19, 9 (2015), 1652–1655.