RFI Mitigation at the MeerKAT site

Anja Schröder
South African Astronomical Observatory
(anja@hartrao.ac.za)

Outline

- RFI team at SKA-SA
- RFI policies
- RFI monitoring on site
- RFI database
- Student projects

RFI team

- RFI team at SKA-SA:
  - Carel van der Merwe
  - now a subsystem equal to others
  - 6 FTE (plus one in the near future)
  - standing contract with 2 consulting groups:
    - MESA (at Stellenbosch)
    - ITC (at Pretoria)
- tasks: checking, testing, policing, monitoring, propagation and attenuation simulation
  - this year: focus on monitoring, both developing and operations

RFI team

- **RFI Working Group:**
  - with representatives from
    - system engineers, science processing team, commissioning team, users, consulting groups, liaison to farmers on site,
  - tasks: advisory; to inform, discuss interdisciplinary problems, combine forces, manage research projects
- **RFI Committee:**
  - 5 people
  - tasks: set policies, make decisions (e.g., on permits)

RFI policies: external

- Astronomy Geographic Advantage Act
- AGA regulations: regulates RFI levels on site
  - -> finalised in ~1 month
  - -> gives operators one year to comply
- e.g.,
  - => GSM at 900MHz
  - => Alkantpan Test Range (military testing site to the north):
    - -> memorandum of agreement on minimising effects
  - => TV (UHF):
    - -> migration from analog to digital, that is, 623MHz; sender been switched off last week

RFI policies: internal

- Policies for permits are now approved and procedures in place
  - every item will be tested:
    - pass -> certificate
    - fail -> note to fix
    - fail but needed -> permit with restrictions in distance from antennae or in time;
- Database with spectrum and temporal behaviour

RFI policies: examples

Examples:

- **X-band receivers failed the test** (strong internal RFI which affect the other receivers on the indexer)
  - **-> permit for 2 weeks for tests on site, then removal**

RFI policies: examples

Examples:

- **S-band receivers failed the test** (strong internal RFI which affect the other receivers on the indexer)
  - **-> permit for 2 weeks for tests on site, then removal**
- **aircraft communication**
  - **cannot deviate flights**
  - **instead restriction to selected few channels**
  - **flight paths predictable -> scheduling**

RFI policies: during construction

- Restrictions increase with project:
  - "what was good enough for KAT 7, is not good enough for MeerKAT"
  - "what was good enough for MeerKAT, is not good enough for SKA Phase 1"

=> standards change: what levels/thresholds to accept

- Audit of MeerKAT site in 2018:
  - not all needs to pass but necessary to understand the RFI situation

-> how best to schedule observations in combination with permits: e.g., day-night

-> compile a 'hitlist' of undesirables

RFI Monitoring

- Real Time Analyser (RTA):
  - omni-directional antenna
  - Version 4: updated with new SKARAB (ROACH3) version to be consistent with SKA backend
    - -> mid-2018
  - now: RTA v3.5 **(within next weeks)**:
    - ROACH2
    - 2 fixed position antennae
    - 1 mobile antenna
    - improved calibration with noise sources
    - remote login functionality
  - monitor occupancy (for permits)
  - monitor compliance with policies (confirm 'switching off')

RFI Monitoring

- **MeerKAT:**
  - 'best' monitor
  - can use receivers on the indexer to observe the horizon (while antanna elevation is 15°)
    - advantage: cooled receiver
  - special monitoring mode of MeerKAT:
    - 20-minute scan of horizon
    - at twilight
    - to verify permits
    - for statistics

RFI Monitoring

- RFI monitoring data (both from MeerKAT and RTA) can be used to test algorithm for flagging, excising etc
  - same archive, same data formats
- customers:
  - RFI team: for verification of internal and external RFI policies
  - students: for R&D
  - scientist: for comparison with their observations
  - MeerKAT telescope: for smart scheduling

RFI classes

- 3 groups:
  - intentional transmitters (external)
    - strong but easy to ID (know frequencies)
    - licensed transmitters
  - unintentional transmitters:
    - usually broadband
    - EMI, e,g., digital with clock, screens, network
    - tricky to deal with
  - transients:
    - short time
    - worst to identify
    - anything with a switch (e.g., air-conditioner)

RFI database

PhD project: Gerald N. Balekaki (UCT Comp Science)

A complete RFI database for the MeerKAT site

- A two-fold record of each RFI:
  - all known culprits (for identification)
  - record of observed RFI on site (for occupancy studies, mitigation, etc)
- Distinguish between known and unknown RFI
  - ⇒ they vary greatly in nature and source
  - ⇒ need to classify
- Determine source and nature of each RFI
- Enable data mining

RFI Database

- Goal: develop prototype to
  - investigate advanced techniques for Big Data
  - discover structure of the data
  - discover nature of the data in order to classify
  - use it

- Key challenges of Big Data: Volume, velocity, variety, veracity (i.e., inherent unreliability of some data)
  - makes it difficult for us to capture, store, analyze and acquire intelligence about them

RFI Database: Design Challenges

Traditional DB are impractical to handle Big Data

- 95% of the entire existing data is unstructured

=>

- New modeling techniques:
  - discover 'nature' of data to classify and fit into the structure of the DB

    -> decision on DB and its purpose
  - extract 'knowledge' (i.e., pattern) from the data

    -> goes into design of DB
- Advanced statistical methods/tools:
  - machine learning

RFI Database: Analytical Challenges

- **Advanced CASE (Computer-Aided Software Engineering) tools needed** (e.g., efficient storage, matching, interfaces):
    - **probabilities instead of facts**
    - **keep only relevant information**

  $\Rightarrow$ **Associative methodology instead of relationships**

RFI Database: Analytical Challenges

Choice of Data Models for large datasets:

Design aspects	Associative DB Models/NoSQL	Relational DB Models
Data Models	No Pre-defined Schema (Highly adaptive)	Defined Schemas
Data Structure	Unstructured Data (makes up 95% of Big Data)	Fairly Structured data
Scaling	Can scale over cheap commodity equipment/servers	Requires bigger, more expensive equipment servers
Development Model	Open source	Closed –Requires licenses, fees etc

RFI Database: Prototype Outcome

- Rapid identification & classification of RFI signals
- Efficient storage of the large amounts of RFI data captured
- Scalability of the prototype such that:
  - it can accommodate other classification modes
  - it can be extended to other observatories

Student projects

- List of past, present and future student projects
  - better organisation of what is needed
  - do not repeat
  - help and supplement each other

$\Rightarrow$ plan to incorporate all RFI endeavours into a webpage or wiki

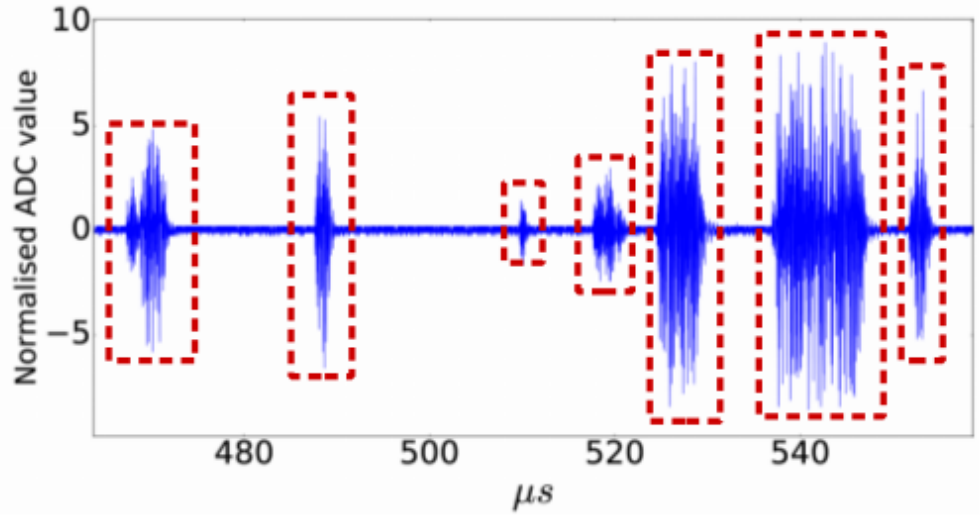
Student projects (e.g.)

- Daniel Czech (PhD at UCT)

  RFI Source Classification

  - Use time domain signals of RFI emitters
  - apply speech recognition techniques

Transients



- Each event = a sequence of tiny interference impulses, or transients.
- Hypothesis:
  - Transients from different events might be similar.
  - All transients might belong to a fixed dictionary.

Student projects (e.g.)

- Daniel Czech (PhD at UCT)

  RFI Source Classification

  - Use time domain signals of RFI emitters
  - apply speech recognition techniques
  - works well on some emitters

Student projects (e.g.)

- Daniel Czech (PhD at UCT)

  RFI Source Classification

  - Use time domain signals of RFI emitters
  - apply speech recognition techniques
  - works well on some emitters
- Kai Staats (MSc at UCT):

  Genetic Programming Applied to RFI Mitigation

  - computer programs are encoded as a set of genes that are modified using an evolutionary algorithm
  - training set from KAT-7

RFI flagging for MeerKAT data

MeerKAT pipeline (for imaging):

- variant of *AOFlagger* (Tom Mauch)
- bootstrapped using a static RFI mask for:
  - ~900 MHz
  - ~1200-1300MHz

Summary

- RFI policies exist and permits are used on anything that goes on site
- RFI monitoring continues and efforts are increased
- New RFI database to store and identify known and new RFI
- RFI flagging in the MeerKAT pipeline based on AOFlagger
- Many student projects covering:
  - RFI control on site
  - RFI monitoring, detection and classification
  - RFI subtraction in the data