Real time data acquisition of commercial microwave link networks for hydrometeorological applications

C. Chwala\textsuperscript{1}, F. Keis\textsuperscript{1}, and H. Kunstmann\textsuperscript{1,2}

\textsuperscript{1}Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research (IMK-IFU), Garmisch-Partenkirchen, Germany
\textsuperscript{2}University of Augsburg, Institute for Geography, Augsburg, Germany

Received: 11 September 2015 – Accepted: 2 November 2015 – Published: 23 November 2015
Correspondence to: C. Chwala (christian.chwala@kit.edu)
Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

The usage of data from commercial microwave link (CML) networks for scientific purposes is becoming increasingly popular, in particular for rain rate estimation. However, data acquisition and availability is still a crucial problem and limits research possibilities. To overcome this issue, we have developed an open source data acquisition system based on the Simple Network Management Protocol (SNMP). It is able to record transmitted- and received signal levels of a large number of CMLs simultaneously with a temporal resolution of up to one second. We operate this system at Ericsson Germany, acquiring data from 450 CMLs with minutely real time transfer to our data base. Our data acquisition system is not limited to a particular CML hardware model or manufacturer, though. We demonstrate this by running the same system for CMLs of a different manufacturer, operated by an alpine skiing resort in Germany. There, the data acquisition is running simultaneously for four CMLs with a temporal resolution of one second. We present an overview of our system, describe the details of the necessary SNMP requests and show results from its operational application.

1 Introduction

Only a decade ago, Messer et al. (2006) introduced the use of commercial microwave link (CML) networks for the quantification of rainfall. Since then, this technique has been applied further in several other countries. Leijnse et al. (2007) performed the first analysis of data from two CMLs in the Netherlands. Increasing the number of CMLs and using the data from countrywide networks Zinevich et al. (2008) and Overeem et al. (2013) were able to derive spatial rainfall information in Israel and the Netherlands respectively. In Germany, Chwala et al. (2012) acquired the first data using data loggers at the CML towers. This limited the number of the CMLs, but provided data with a very high power resolution.
Just recently, Doumounia et al. (2014) used CML data from a cell phone provider in Burkina Faso and showed that this technique is also feasible in West Africa. Due to the very coarse station network and the absence of weather radars, the additional CML derived precipitation information is particularly important in African countries. Fortunately, the growing demand for mobile communication ensures an increasing number of CMLs also in this region.

Besides its application for the quantification of rainfall, CML data can also be used for the estimation of atmospheric humidity (David et al., 2009).

However, the availability of CML data remains one of the crucial hurdles for research and future operational applications at national meteorologic services and water authorities.

The software we have developed and which is running operationally in Germany, strives to improve this situation. It is capable of acquiring data of homogeneous or heterogeneous CML networks at high temporal resolution in near real time. In this article we describe the technical background, the details of the software and show results from its applications. Section 2 describes the problem of limited CML data availability for research. Section 3 introduces the concept of acquiring CML data via SNMP. Section 4 explains the details of the implementation of our real-time data acquisition system. Section 5 gives examples from two operational applications of the system. And Sect. 6 discusses the transferability to other CML networks.

With this development and initiative we finally aim not only to provide a useful tool for the hydrometeorological research with CMLs, but also to encourage other CML operators to open their networks for custom data acquisition.
2 The problem: limited availability of CML data

2.1 Necessary data for rain rate estimation with CMLs

To be able to estimate rain rates with a microwave link, information of the attenuation along the path is necessary. This attenuation is strongly correlated with the rain rate along the path. To determine the attenuation, the transmitted (TX-level) and the received power (RX-level) of the links have to be known. For older CML hardware, the TX-level is typically constant, but modern systems support an adjustment of the TX-level via Automatic Transmission Power Control (ATPC) to compensate attenuation events and keep the RX-level in the receiver’s optimal range. Hence, recordings of the RX-, and very often also TX-levels are necessary, ideally at a time resolution of minutes to capture the temporal variability of rainfall.

2.2 Situation for the CML operators

One of the biggest advantages of the hydrometeorological usage of CML networks is that they already exist, are well maintained and provide a good spatial coverage in inhabited areas (Overeem et al., 2013). The large country-wide CML networks are operated by cell phone providers, or by CML manufacturers for a cell phone provider. Smaller networks are operated locally, e.g. by companies to establish wide area networks (WAN) between different remote sites.

For the operators of CML networks, the attenuation that is caused by rainfall is only unwanted disturbance, since it perturbs the communication along the CML in case of strong attenuation events. In periods with extremely heavy rainfall this may even lead to total loss of connection.

Strong attenuation events are not the only cause of failures in a CML network, though. Hardware and software problems occur and have to be fixed promptly to assure constant high availability of a network’s full functionality. Hence, the CML networks...
are continuously monitored. But for the monitoring, neither the exact point in time, nor the exact amount of precipitation that may have caused a system failure is important.

As consequence the network operators store TX- and RX-level data only in a coarse resolution (15 min, hourly, daily) or do not store that data at all.

2.3 State of the art CML data acquisition for research

For the purpose of monitoring, modern CML systems support the storage of a so called performance report, which typically contain the TX-level and RX-level. Depending on the hardware type and the chosen settings these reports are generated every 15 min, every hour or daily. The most common setting is a temporal resolution of 15 min and the recording of the minimum and maximum values within this period. The TX- and RX-level resolution is typically 0.3 or 1 dB, depending on the hardware. This performance data is stored e.g. by T-mobile in the Netherlands (Overeem et al., 2013) and by cell phone providers in Israel (Zinevich et al., 2008) and is made available for research. Data transfer from the providers to the researches is usually made in chunks on daily or irregular basis.

In Germany the situation with our cooperation partner Ericsson is different. Ericsson operates the CML network for T-mobile, but does not store the CML performance reports. Hence, in our initial approach, we used small data loggers directly at the CML towers to record the RX-level, which is available at the analog output of the automatic gain control (AGC) of older hardware (Chwala et al., 2012). The TX-level was constant at the CMLs which we equipped. Data was sent to a database from the data loggers via GSM every day. This approach had the advantage of a very high power resolution of the RX-level, about 0.03 dB, only limited by the resolution of the module’s analog to digital converter. However, it also had one main drawback. Equipping one link did not only involve the costs for the data acquisition module (approximately EUR 100), but also the expenses for an Ericsson technician who did the actual installation at the towers. Hence, using dedicated data acquisition modules at the towers was not suitable for an extension to a country-wide coverage, involving several thousands CMLs.
In joint cooperation with Ericsson Germany we thus developed a solution which can easily be scaled up for a larger number of CMLs and which at the same time improves the temporal resolution compared to the available data from the performance reports.

3 The solution: custom data acquisition via SNMP

3.1 The simple network management protocol (SNMP)

Since CMLs are built to form communication networks, it is easy to communicate with an individual CML when having access to the network. Each CML is accessible via its IP address, which is typically from a private IP address space. For service and configuration related communication the Simple Network Management Protocol (SNMP) is used, which was developed to allow standardized communication between network attached devices, i.e. routers, printers or computers.

A SNMP request is defined by a target IP address and the SNMP Object Identifier (OID). The OID defines the actual command that has to be evoked. The OIDs are defined in the Management Information Base (MIB). There are general MIBs which define OIDs that every system supports, like `sysUpTime`, which would request the uptime of the system. For additional functionality the MIBs can be extended. E.g. `genEquipRfuStatusTemperature` is the OID to query the hardware temperature of the radio unit of a Ceragon IP10 CML.

3.2 SNMP for monitoring a CML

In large CML networks, like the ones operated by the cell phone providers, SNMP is used to remotely configure the CMLs and continuously monitor them for failures. Querying the TX- and RX-level is a typical functionality of the monitoring capabilities and is, to our knowledge, supported on all modern CMLs.

The OIDs to request the TX- and RX-level are not standardized and differ from CML hardware to CML hardware, even for the same manufacturer. E.g., `xfRFCurrentInput-`
Power is the base OID for requests of the current output power for a Ericsson MINK-LINK Traffic Node CML while \textit{gnOduStatusXReceiveLevel} is the one for a Ceragon IPMax and \textit{genEquipRfuStatusRxLevel} for a Ceragon IP10 system. For each system they have to be looked up in the MIBs which are available to the CML owners from the manufacturers.

Since modern CML systems like the Ericsson MINK-LINK Traffic Nodes support the operation of several CMLs managed by one main control unit, the derivation of the correct OID requires a further step, though. The base OID has to be extended to specify which sub-system, hence which CML, has to be accessed. Furthermore, some modern systems make it possible to query data not only for the near end of a CML, to which the SNMP requests are sent, but also for the opposite site, the far end. This also has to be encoded in the OID.

Appendix A explains in detail how the correct OIDs for different systems are built.

3.3 Concept of a SNMP based real time data acquisition

Based on the knowledge of the SNMP communication capabilities of the CMLs and respecting the security requirements of CML operators, the concept of our custom CML data acquisition was built on the following findings:

- Necessary CML data can be queried via SNMP.
- SNMP request can be performed simultaneously for different host systems.
- SNMP access is only possible from within the private network of the CML operators.
- Access to the private CML operator access will not be granted from outside because of security concerns.

After initial testing of a proof of concept and further negotiations with Ericsson Germany the resulting concept, shown in Fig. 1, consisted of:
– A data acquisition server located inside the private subnet of the CML operators running a custom software that performs the SNMP requests.

– The custom data acquisition software which is capable of performing simultaneous SNMP request to a large number of CMLs.

– A solution to instantaneously move the data out of the private subnet to a data base server running at our institute.

While the concept looks simple, its realization is challenging. The current operational status of the data acquisition is not only a technical, but also a superordinate success, since the CML operators had to grant access to their highly secured private network to allow the operation of the custom data acquisition software. Trust is an important prerequisite, for which thorough testing of the software was necessary. We were able to do the necessary tests during development in a non-operational environment with several CMLs at an Ericsson test facility.

4 The implementation: pySNMPdaq

4.1 Overview

Our open source real-time SNMP based CML data acquisition software is written in Python and hence was given the name pySNMPdaq.

To assure a constant timing of the simultaneous SNMP request, pySNMPdaq is divided into three types of sub-processes using the Python standard multiprocessing library:

1. Timer: The timer Process which continuously triggers events at a fixed temporal resolution. The smallest time step is one second, with an accuracy of 10 ms. It avoids the typical drift of a simple while-sleep loop construct by syncing to the hardware clock of the system it is running on.
2. **snmpDAQSession**: The data acquisition process, one for each CML, which handles the SNMP communication with the CMLs. The OIDs which are used for the SNMP requests are defined in a configuration file and can be different for each snmpDAQSession. To assure continuous real-time operation, the SNMP communication is asynchronous and avoids blocking due to failing request which have not timed out. The SNMP requests are triggered via a Queue by the timer process. The queried data is put into the Queue of the data handler process.

3. **DataHandler**: The process to which all queried data is fed via a Queue and which writes the data to file or sends it to an external server via scp (Secure copy). Writing, closing and transferring files is triggered by the timer via a Queue.

Figure 2 shows a schematic of the interplay of the three different process types.

### 4.2 Requirements

The main requirements, besides software dependencies, to successfully run `pySNMP-daq` are:

- Access to a computer with SNMP connectivity to the CMLs.
- Knowledge of the CML IP addresses.
- Knowledge of the CML hardware to assure the usage of the correct OIDs for each CML.

To be able to run `pySNMPdaq` the following Python packages have to be available:

- `pySNMPdaq` ([http://github.com/cchwala/pySNMPdaq](http://github.com/cchwala/pySNMPdaq))
- `pandas` version $\geq 0.14$
- `numpy` version $\geq 1.8$
- `netsnmp` (Python bindings for the net-snmp C-library).
4.3 Application

When the above requirements are met, a list of IP addresses, each with a set of OIDs, has to be compiled and the global configuration file has to be adjusted to define sampling intervals and data handling. pySNMPdaq can then be started and will be running as a background process.

When run in an operational CML network, care should be taken, to not interfere with its normal operation. Hence a slow increase of the number of CMLs and a reasonable sampling frequency should be considered.

Our test runs at the Ericsson test facility and additional calculation of the caused traffic, show that a continuous data acquisition of several hundred CMLs at a 1 s interval does not impair the network. This has not been tested operationally in a large network, though. A sampling interval of one minute was agreed upon with Ericsson for the data acquisition in their operational CMLs.

5 Example data

5.1 Data from CMLs for cell phone backhaul network

In summer 2014 we started to run pySNMPdaq operationally on a server in the CML network that Ericsson Germany operates as backhaul for the T-mobile cell phone net. Slowly increasing the number of simultaneous queried CMLs, we are currently receiving data from 450 Ericsson MINK-LINK TN CMLs every minute with a time delay of only several seconds.

Figure 3 shows a record of one CML over the course of three days. The CML has a length of 7.9 km and uses frequencies of 22.078 and 23.086 GHz for the two directions. Plotted is the difference between the transmitted and received level (TX-RX level). Quantization is 1 dB for the TX-level and 0.3 dB for the RX-level. The CML is configured as a 1+1 protection system using a second transmit and receive unit at the
near end as a Hot Standby. This configuration is often used at sites where the outdoor units, which contain the transmitter and receiver, are mounted in places which are difficult to reach. There, an exchange of an outdoor unit is often more expensive than the unit itself.

The lower part of Fig. 3 is a zoomed-in version showing one of the strong attenuation events in more detail. One can see that the three channels (far–near, near–far, far–near) agree very well. Interestingly there are, however, small differences even between far–near and far–near, which operate in the same direction and share the transmitter at the far end. The differences hence must be caused by small variations in the two receivers and the subsequent rounding to the quantized values that are available via SNMP.

5.2 Data from CMLs for skiing resort intranet

Besides the operation for the CMLs in the Ericsson cell phone backhaul network, we also installed pySNMPdaq at the Bayerische Zugspitzbahn, the operator of the local alpine skiing resort. They connect the summit stations, e.g. on Germany’s highest mountain, the Zugspitze, with the valley via Ceragon CMLs. Since their network only comprises four CMLs, computational demand was low and pySNMPdaq was run on a low performance netbook placed at the IT-center of the Bayerische Zugspitzbahn. Since the operations of this CML network is not as crucial as the one that serves as the cell phone network backhaul, we were allowed to run the data acquisition with one second temporal resolution.

Figure 4 gives an example of TX-RX level records. Note that the Ceragon CMLs have 1 dB quantization of the received signal level. The transmit level was constant. In the zoomed-in plot the high temporal resolution is revealed. Since data at 1 s intervals is not necessary for rain rate estimation, the quantization could be compensated by averaging. In situations with low fluctuations of the signal level, this however leads to artifacts.
The operation at the Bayerische Zugspitzbahn also proofed that *pySNMPdaq* is capable of acquiring data from different types of CML hardware, since three of the four Ceragon CMLs are IP10 models and one is a IPMax system.

### 6 Transferability to other CML networks and CML hardware

As has been mentioned in Sect. 5.2, *pySNMPdaq* can acquire data for CML networks with heterogeneous hardware as long as the CML hardware supports querying the received, and if necessary the transmitted signal level. To our knowledge this is true for most, if not all, modern CML systems. The only challenge is to derive the correct OIDs for each hardware type.

Given its flexibility regarding CML hardware, *pySNMPdaq* should be usable in all CML networks worldwide.

### 7 Conclusions

We have introduced our newly developed operational real time data acquisition system for CML data, which is available as open source software at http://github.com/cchwala/pySNMPdaq. This system is the basis for continuous countrywide access to CML data for line integrated precipitation quantification.

Its application is the first of this kind, where a specifically design software for scientific purpose is run in a large operational CML network. The software is capable of acquiring data from different CML hardware models and manufacturers and, hence, is applicable in any CML network.

We have shown that it is possible to provide CML data with high temporal resolution (one minute or better) and in near real-time (only several seconds delay).

In particular the almost instantaneous availability is a crucial prerequisite for future operational applications of CML data for quantitative precipitation estimation (QPE). Retrospective and theoretical studies have already shown the potential of CMLs to...
improve QPE from weather radars (Trömel et al., 2014) and to considerably increase the accuracy of urban drainage modeling (Fencl et al., 2013). But the CML data must be available at the time the rain is falling to be able to predict and protect. This is now possible.

For Germany, the next step is a further extension of the data acquisition system to provide a countrywide coverage using several thousand CMLs. Furthermore, negotiations have been initialized with cell phone providers in Palestine and Burkina Faso (Gosset et al., 2015) to improve the data acquisition of CMLs for rain rate estimation in these data scarce regions.

Appendix A: Example OIDs for received signal level

This appendix describes the OIDs for requesting the received signal level for the CML hardware that was used to record the data shown in Sect. 5.

Ericsson MINI-LINK Traffic Node systems support the operation of several CMLs as subsystems of one main control unit. The different CML subsystems are defined by their slot number. On top of that it has to be defined whether the near end or the far end should be queried. Building on the base OIDs for requesting the RX-level, 1.3.6.1.4.1.193.81.3.4.3.1.3.1.10, the additional information is encoded in the so called interface descriptor. As already observed by Wang et al. (2012), the interface descriptor is a cryptic number, e.g. 2146697473 for slot number 2 at the near end. The logic behind the cryptic number is revealed when different interface descriptors are given in hexadecimal representation. Table A1 lists OIDs and interface descriptors for several Ericsson MINK-LINK Traffic Node subsystems. Going from slot 2 at the near end to slot 3 at the near end the interface descriptor changes from 0x7ff40101 to 0x7ff40181. This is equal to a switch of bit number 8 from 0 to 1. Similarly the difference between near end and far end is a switch of bit number 19 from 1 to 0.

Other CML hardware uses different OIDs and different methods to distinguish the subsystems. For example a Ceragon IPMax uses $\text{gnOduStatusXReceiveLevel}$ as base
OID to query the current received signal level. It supports two subsystems, in drawer1 and drawer2, for which the base OID has to be altered. Table A1 gives these OIDs together with that of the older Ceragon IP10 system which does not support subsystem.

Acknowledgements. We want to gratefully thank our cooperation partners at Ericsson Germany: Jürgen Fritz for the support with our first data acquisition based on the data loggers and for introducing us to the right people to make the move to SNMP possible. Jürgen Alschbach for the technical support while figuring out the details of the CML SNMP connectivity in the test lab. Ronald Riedel for supporting our ideas and thus for making it possible to move from the lab to the operational network. Reinhard Gerigk and Declan Forde for the initial and continuous support with the actual operational system.

Furthermore we want to thank Christian Scharpf from Bayerische Zugspitzbahn AG for making it possible to record CML data in the alpine terrain in our backyard.

This work was supported by the German research foundation (DFG) through the project Integrating Microwave Link Data For Analysis of Precipitation in Complex Terrain: Theoretical Aspects and Hydrometeorological Applications (IMAP). Further support was provided by the Stiftung Energieforschung Baden-Württemberg. We acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of Karlsruhe Institute of Technology.

The article processing charges for this open-access publication were covered by a Research Centre of the Helmholtz Association.

References


Table A1. Variations of the OIDs for requesting the current received signal level for different CML systems and, if applicable, for their subsystems. The Ericsson MINI-LINK Traffic Node (TN) uses an interface descriptor to identify the subsystems and can is able to also access data from the far end of the CML. The Ceragon IPMax systems support only two subsystem is separate drawers, while the Ceragon IP10 does not support subsystems.

<table>
<thead>
<tr>
<th>Hardware</th>
<th>OID (as text)</th>
<th>OID and interface descriptor (as number)</th>
<th>Interface descriptor (as hex)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ericsson MINK-LINK TN</td>
<td>xIRFCurrentInputPower (slot 2 near) 1.3.6.1.4.1.193.81.3.4.3.1.10.2146697473</td>
<td>0x7ff40101</td>
<td></td>
</tr>
<tr>
<td></td>
<td>xIRFCurrentInputPower (slot 2 far) 1.3.6.1.4.1.193.81.3.4.3.1.10.2146697473</td>
<td>0x7ef40101</td>
<td></td>
</tr>
<tr>
<td></td>
<td>xIRFCurrentInputPower (slot 3 near) 1.3.6.1.4.1.193.81.3.4.3.1.10.2146697601</td>
<td>0x7ff40181</td>
<td></td>
</tr>
<tr>
<td></td>
<td>xIRFCurrentInputPower (slot 3 far) 1.3.6.1.4.1.193.81.3.4.3.1.10.2146697601</td>
<td>0x7ef40181</td>
<td></td>
</tr>
<tr>
<td></td>
<td>xIRFCurrentInputPower (slot 4 near) 1.3.6.1.4.1.193.81.3.4.3.1.10.2146697729</td>
<td>0x7ff40201</td>
<td></td>
</tr>
<tr>
<td></td>
<td>xIRFCurrentInputPower (slot 4 far) 1.3.6.1.4.1.193.81.3.4.3.1.10.2146697729</td>
<td>0x7ef40201</td>
<td></td>
</tr>
<tr>
<td>Ceragon IPMax</td>
<td>gnOduStatusXReceiveLevel.drawer1 1.3.6.1.4.1.2281.3.1.5.1.5.3</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td></td>
<td>gnOduStatusXReceiveLevel.drawer2 1.3.6.1.4.1.2281.3.1.5.1.5.4</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Ceragon IP10</td>
<td>genEquipRfuStatusRxLevel 1.3.6.1.4.1.2281.10.5.1.1.2.1</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. Overview of the basic concept of our real-time CML data acquisition system at Ericsson. The CMLs are shown together with one main control unit per site (this is a simplified example) which can be identified via its private IP address. The DAQ server inside the private subnet requests the CML signal levels via SNMP and sends the data immediately to a database server at our institute (KIT).
**Figure 2.** Schematic of the structure of our real time data acquisition software *pySNMPdaq* with its three different kind of processes. The inter-process communication is established via Queues. With the SNMP requests the data flows from the CMLs to the *snmpDAQSessions* and from there to the *DataHandler* which writes data to files in regular intervals and can transfer the data to a further server.
Figure 3. Example data record of an Ericsson MINK-LINK TN CML used in the backhaul of a cell phone network. The CML has a length of 7.9 km and uses frequencies of 22.078 and 23.086 GHz for the two directions. The direction from the far end to the near end uses a second receiver as Hot standby 1+1 system. (top) The solid lines show the 15 min averages of the TX-RX level (transmitted minus received signal level). The lines with the lighter color show the raw data with one minute resolution. (bottom) Zoomed in version of the plot above showing only the raw 1 min data.
Figure 4. Example data record of a Ceragon IPMax CML used for the intranet connection between a summit and a valley stations of an alpine skiing resort. The CML has a length of 4.7 km and uses frequencies of 22.022 and 23.033 GHz for the two directions. (top) The solid lines show the 15 min averages of the TX-RX level (transmitted minus received signal level). The lines with the lighter color show the raw data with one second resolution. (bottom) Zoomed in version of the plot above showing only the raw 1 s data.