Making in-network data processing decisions based on pragmatic value of information

Lotzi Bölöni

Dept. of Electric Engineering and Computer Science
University of Central Florida- Orlando, FL

June 11, 2012

1 Introduction

Sensor networks

- Sensor networks:

Sensor networks

- Sensor networks:
    - distributed systems

Sensor networks

- Sensor networks:
  - distributed systems
  - nodes embedded in an environment

Sensor networks

- Sensor networks:
  - distributed systems
  - nodes embedded in an environment
  - collect information through their sensors

Sensor networks

- Sensor networks:
  - distributed systems
  - nodes embedded in an environment
  - collect information through their sensors
  - transmit the information to customers

Sensor networks

- Sensor networks:
  - distributed systems
  - nodes embedded in an environment
  - collect information through their sensors
  - transmit the information to customers
- Overall objective

Sensor networks

- Sensor networks:
  - distributed systems
  - nodes embedded in an environment
  - collect information through their sensors
  - transmit the information to customers
- Overall objective
  - Minimize consumption of a set of scarce resources

Sensor networks

- Sensor networks:
  - distributed systems
  - nodes embedded in an environment
  - collect information through their sensors
  - transmit the information to customers
- Overall objective
  - Minimize consumption of a set of scarce resources
  - Maximize some metric of sensing quality

Scarce resources

- Energy
- Network bandwidth
  - For real time communication
- Data muling delay
  - How soon will the data mule come by?
- Human resources
  - Attention span and cognitive load of the customer

Some thoughts about the balance between networking and sensing

- For some sensor types, the amount of data collected is trivially small
  - Eg. most types of thermometers
- For others, it can be huge
  - Eg. imaging

When real time networking fails

- A 100 euro video camera can generate 28Mbps (1080p/60)
  - 100 such cameras exceed just about any networking technology, except dedicated optical fiber to every camera.
- What to do?
- **Data muling**: put a 64GB SD card in each sensor
  - physically collect it every 5hrs (about 60GB will be filled)
  - long delay in obtaining the data
- **Local rating**
  - allow the local not do decide whether the data is important or not
  - prioritize high-value data (either on fast networking, or by bringing over the data mule)
  - must put processing intelligence in the local node
- **Local summarization**
  - Create summaries / digests in the local node
  - Prioritize the transmission of digests
  - Still requires local intelligence
  - The value of the data can be determined in collaboration

Measuring sensing quality: quantity, accuracy, pragmatics

- Information theoretic metrics of information
  - measure the **quantity** of information, without regards to its meaning
  - a report take the same amount of bits whether correct and incorrect

Measuring sensing quality: quantity, accuracy, pragmatics

- Information theoretic metrics of information
  - measure the **quantity** of information, without regards to its meaning
  - a report take the same amount of bits whether correct and incorrect
- Accurracy metrics
  - measure the correspondence between the information and reality
  - try to lower the error, the delay of obtaining the information
  - quantity of information is an upper limit on accuracy
  - do not consider the use to which the information will be put

Measuring sensing quality: quantity, accuracy, pragmatics

- Information theoretic metrics of information
  - measure the **quantity** of information, without regards to its meaning
  - a report take the same amount of bits whether correct and incorrect
- Accurracy metrics
  - measure the correspondence between the information and reality
  - try to lower the error, the delay of obtaining the information
  - quantity of information is an upper limit on accuracy
  - do not consider the use to which the information will be put
    - the accuracy is the same whether we are tracking a deer or a terrorist

Measuring sensing quality: quantity, accuracy, pragmatics

- Information theoretic metrics of information
  - measure the **quantity** of information, without regards to its meaning
  - a report take the same amount of bits whether correct and incorrect
- Accurracy metrics
  - measure the correspondence between the information and reality
  - try to lower the error, the delay of obtaining the information
  - quantity of information is an upper limit on accuracy
  - do not consider the use to which the information will be put
    - the accuracy is the same whether we are tracking a deer or a terrorist
- Pragmatic metrics
  - measure the utility of the information to the customer - thus it depends on the **actions** taken by the customer
  - accuracy of information is an upper limit on pragmatic utility

Pragmatic value of information

- Probably useful to express it in form of **money**
- Must consider
  - Content - what aspect of reality the information refers to
  - Accuracy - correspondence between reality and information
  - Latency - time between evaluation and the events the information refers to
  - Novelty - did I know this before?

Decisions of the customer

- Pragmatic value depends on the decisions of the customer
- **Disruptive decisions**
  - change the algorithm for calculating the pragmatic value of information.
  - usually: commitment to a new plan - e.g. raising an alarm.

Decisions of the customer

- Pragmatic value depends on the decisions of the customer
- **Disruptive decisions**
  - change the algorithm for calculating the pragmatic value of information.
  - usually: commitment to a new plan - e.g. raising an alarm.
- **Incremental decisions**
  - do not change the algorithm for calculating the pragmatic value of information

Why considering pragmatic value is useful

- Prioritize information which is acted on in real time
- Buffer information which will not be acted on real time
- Discard worthless information
- Target the amount of information to the action being taken

Why considering pragmatic value is challenging

- The value of information changes in time (even retroactively)

Why considering pragmatic value is challenging

- The value of information changes in time (even retroactively)
- The system must support the decision making process of the human or agent decision makers
  - thus it must be aware of the impact of various pieces of information on the decisions.

Why considering pragmatic value is challenging

- The value of information changes in time (even retroactively)
- The system must support the decision making process of the human or agent decision makers
  - thus it must be aware of the impact of various pieces of information on the decisions.
- The system must support recovery from erroneous costumer decisions.

Customers building a model of the world

- Customer C at time t
- Collection of *raw data* $D = \{d_1, d_2, \ldots d_k\}$
- Build a *model of the world* $M = f_M(D, t)$

Customers building a model of the world

- Customer C at time t
- Collection of *raw data* $D = \{d_1, d_2, \ldots d_k\}$
- Build a *model of the world* $M = f_M(D, t)$

Nature of the model

We will treat the model as a black box. The actual **data structure** can range from a single scalar to lists of values, raster-based environmental models, linear or non-linear predictors, confidence ranges and other models of arbitrary complexity. The **modeling function** can use a combination of techniques such as filtering, interpolating, extrapolating, system identification and others.

The value of the model

- The *value* of the model is a scalar:
  $V = f_V(M) = f_V(f_M(D, t)) \in \mathbb{R}$
- Scores the detail, accuracy and timeliness of the model, weighted by the interests of the customer.

The value of an information chunk d

- sensor node has a chunk of data d
- if sent to customer, added to raw data $D' = D \cup d$
- updated model $M' = f_M(D', t)$
- value of data chunk $V(d, t) = f_V(M') - f_V(M) \geq 0$

The value of an information chunk d

- sensor node has a chunk of data d
- if sent to customer, added to raw data $D' = D \cup d$
- updated model $M' = f_M(D', t)$
- value of data chunk $V(d,t) = f_V(M') - f_V(M) \geq 0$

Worthless data

If $V(d,t) = 0$, the data was worthless for the customer. This can happen if:

- the data was already received from other sources
- the data is of no interest to the customer

How to build a value function?

- Common sense:
  - more accurate models are better
  - if an information chunk does not change the model, its value is zero.
- Many functions conform to this!
- Systematic approach in terms of **pragmatics**

Pragmatic value of information

- Let us denote with $f_V(M, A)$ the pragmatic value of a model of the world M for an agent A. We define this value to be the cummulative value of the actions which had been taken while using the model as source of information.
- Are we just pushing the problem of defining value further?

Pragmatic value of information

- Let us denote with $f_V(M, A)$ the pragmatic value of a model of the world M for an agent A. We define this value to be the cummulative value of the actions which had been taken while using the model as source of information.
- Are we just pushing the problem of defining value further?
- Yes, but...

Pragmatic value of information

- Let us denote with $f_V(M, A)$ the pragmatic value of a model of the world M for an agent A. We define this value to be the cummulative value of the actions which had been taken while using the model as source of information.
- Are we just pushing the problem of defining value further?
- Yes, but...
- It is easier to define values in the application domain!

Pragmatic value of information

- Let us denote with $f_V(M, A)$ the pragmatic value of a model of the world M for an agent A. We define this value to be the cummulative value of the actions which had been taken while using the model as source of information.
- Are we just pushing the problem of defining value further?
- Yes, but...
- It is easier to define values in the application domain!
  - Many terminal states have natural values

Pragmatic value of information

- Let us denote with $f_V(M, A)$ the pragmatic value of a model of the world M for an agent A. We define this value to be the cummulative value of the actions which had been taken while using the model as source of information.
- Are we just pushing the problem of defining value further?
- Yes, but...
- It is easier to define values in the application domain!
    - Many terminal states have natural values
    - Contrast this with trying to assign a value to 100kB of data!

The sources of decisions

- The main sources of decisions is (surprise!) the sensor readings themselves
- The local node:
  - Has access to the full local data in real time
  - Limited processing power
- The central node:
  - Has access to the global picture
  - More processing power
  - Human decision makers might be present
  - Delayed or incomplete access to the local data

The problem of being wrong

- Sensors are noisy, processing algorithms imperfect, humans inattentive:
  - Errors happen!
- The judgements will be inevitably fuzzy/probabilistic/degree based.
  - But decisions are usually binary: transmit / not transmit, send the DM or not.
- Underestimating and overestimating the value of information can both be wrong
  - Having a fixed *safety margin* doesn't buy us anything when we are resource constrained - it won't change the resource allocation
- One solution: estimate uncertainty in judgement

Dempster-Shafer theory of evidence

- Dempster-Shafer mathematical theory of evidence
  - find a belief function which combines evidence from various sources
  - can be seen as a generalization of probability
- Uses two values: **belief** and **plausibility**
  - ... instead of a single probability value
  - belief $<$ plausibility
  - mass function: assigns weights to the powerset of possible outcomes
  - belief $=$ mass of evidence for
  - plausibility $=$ 1 - mass of evidence against

Using D-S for the pragmatic value of information

- The **belief** is used as an input to the operator in order to support a disruptive decision.
  - a belief > 0.5 leads to a classification decision.
  - A human operator can override this value.
- The **plausibility** used to assess the likelihood of being wrong
  - eg. what I see is probably just a scratch, but it is still plausible that it is a crack.
  - the system must prepare for the possibility of the customer changing his mind about the classification.

Conclusions - the big picture

- Whenever in a sensor network we have scarce resources, we need to decide what data to send now, what to delay, what to discard
- The ultimate use to which the data can be put (the pragmatic value) is a natural basis to answer such questions
- Value of information resides in:
  - deciding what actions to take
  - providing data for successful completion of the actions
  - providing support for correcting bad decisions
- All these are application dependent, but it is possible to establish a general purpose theoretical framework

Conclusions - the big picture

- Whenever in a sensor network we have scarce resources, we need to decide what data to send now, what to delay, what to discard
- The ultimate use to which the data can be put (the pragmatic value) is a natural basis to answer such questions
- Value of information resides in:
  - deciding what actions to take
  - providing data for successful completion of the actions
  - providing support for correcting bad decisions
- All these are application dependent, but it is possible to establish a general purpose theoretical framework
- ... there is much left to do

Questions?