Scalable Meter Data Management in Electric Power Grids

Vinod Namboodiri

Wichita State University

(vinod.namboodiri@wichita.edu)

PSERC Webinar

October 6, 2015

Acknowledgements

- Students: Babak Karimi, Vishnu Cherusola
- Collaborators: Murtuza Jadliwala, Visvakumar Aravinthan from Wichita State, and Anurag Srivastava from Washington State
- Industry advisors of PSERC project S-54.
- PSERC for support for S-54: Towards a Privacy-Aware Information-Sharing Framework for Advanced Metering Infrastructures

Presentation Outline

- Data volume as an issue
- Overview of solution approaches
- Data concentration for AMI
- Application-quality aware data aggregation for AMI
- Future work

Data and Smart Grids

- Collect, store and manage huge quantities of data -- from millions of smart meters

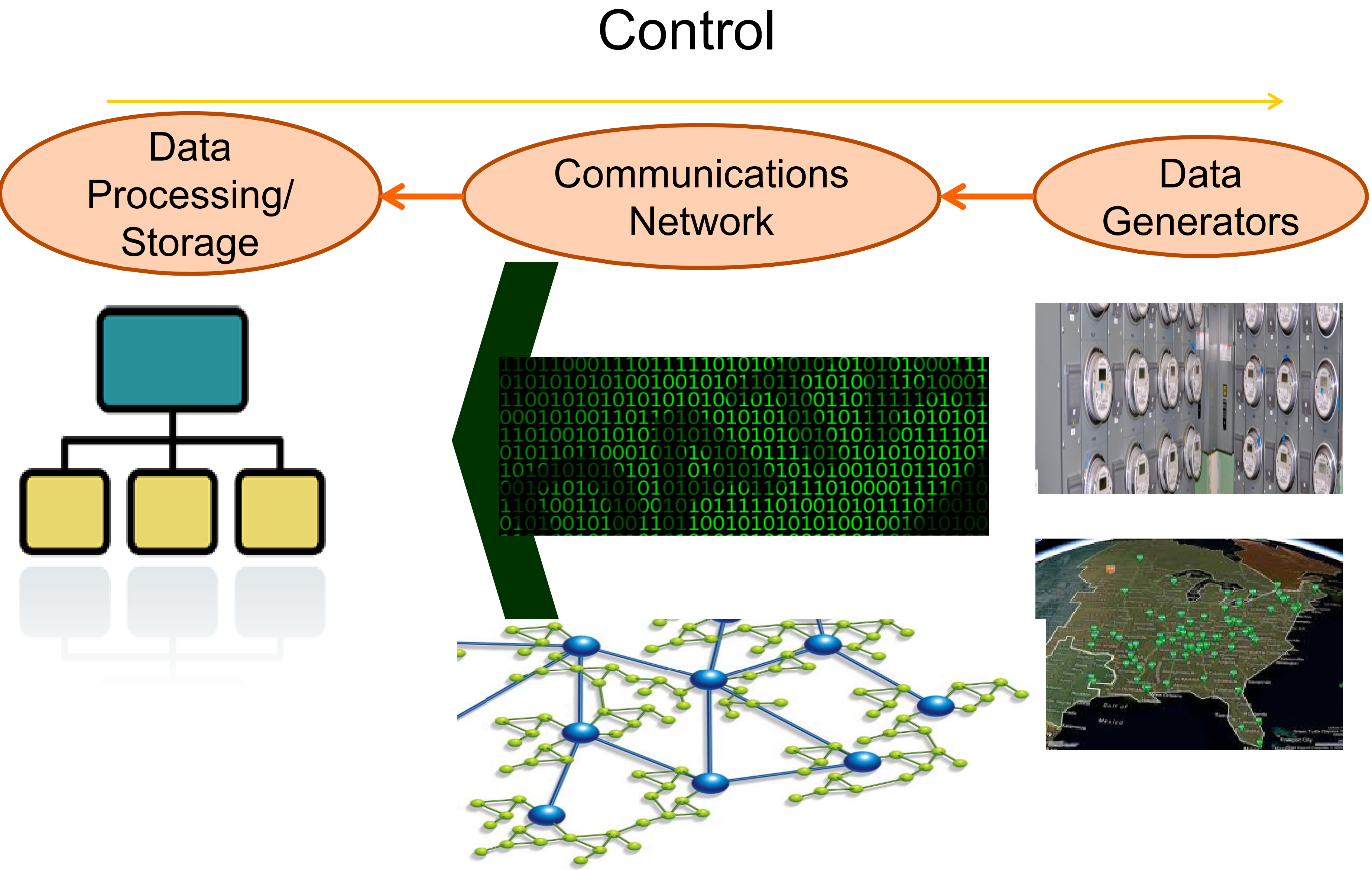
  - Advanced Metering Infrastructure (AMI)
  - A popular tool to modernize the electric grid (efficiency, renewable integration)
  - Facilitates two-way communication between smart meter and control center
  - Captures and transmits energy-use on a hourly or sub-hourly basis
  - As opposed to the current meters that provide daily energy usage and monthly bill
  - Expected to facilitate consumer participation in the Smart Grid

Data and Smart Grids

- Collect, store and manage huge quantities of data -- high-frequency samples from phasor measurement units (PMUs)

- Wide-area, Monitoring, Protection, and Control (WAMPAC)
- Synchrophasors or Phasor Measurement Units (PMUs) take 60 Hz snapshots of the state of the grid and send over the communications network

Data Stream – An Abstract View



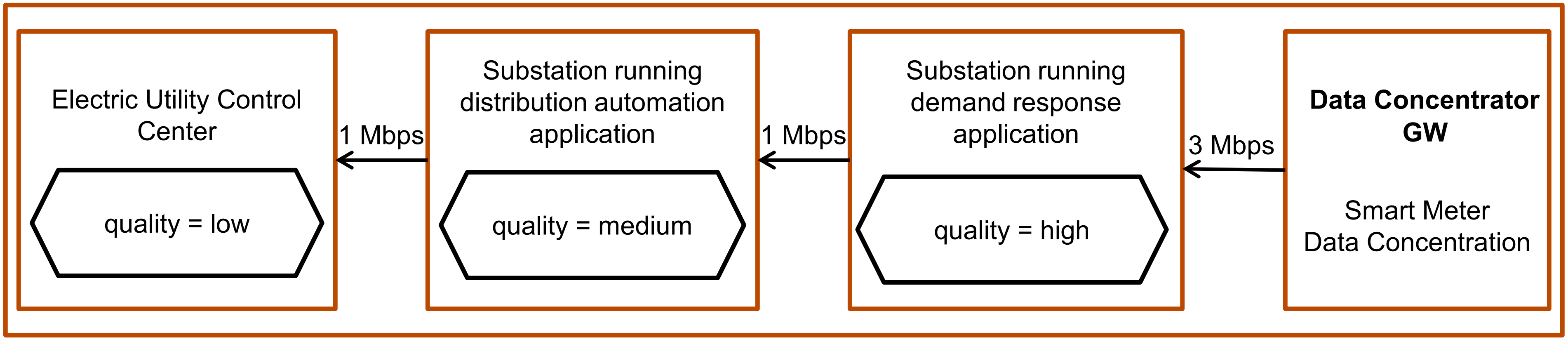
Why is Data Volume an Issue?

- Scaled up to many thousands of meters or synchrophasors sending periodic data, the volume of data can be a burden
  - Limited-bandwidth last-mile networks
  - Perhaps even finer granularity data collection in future

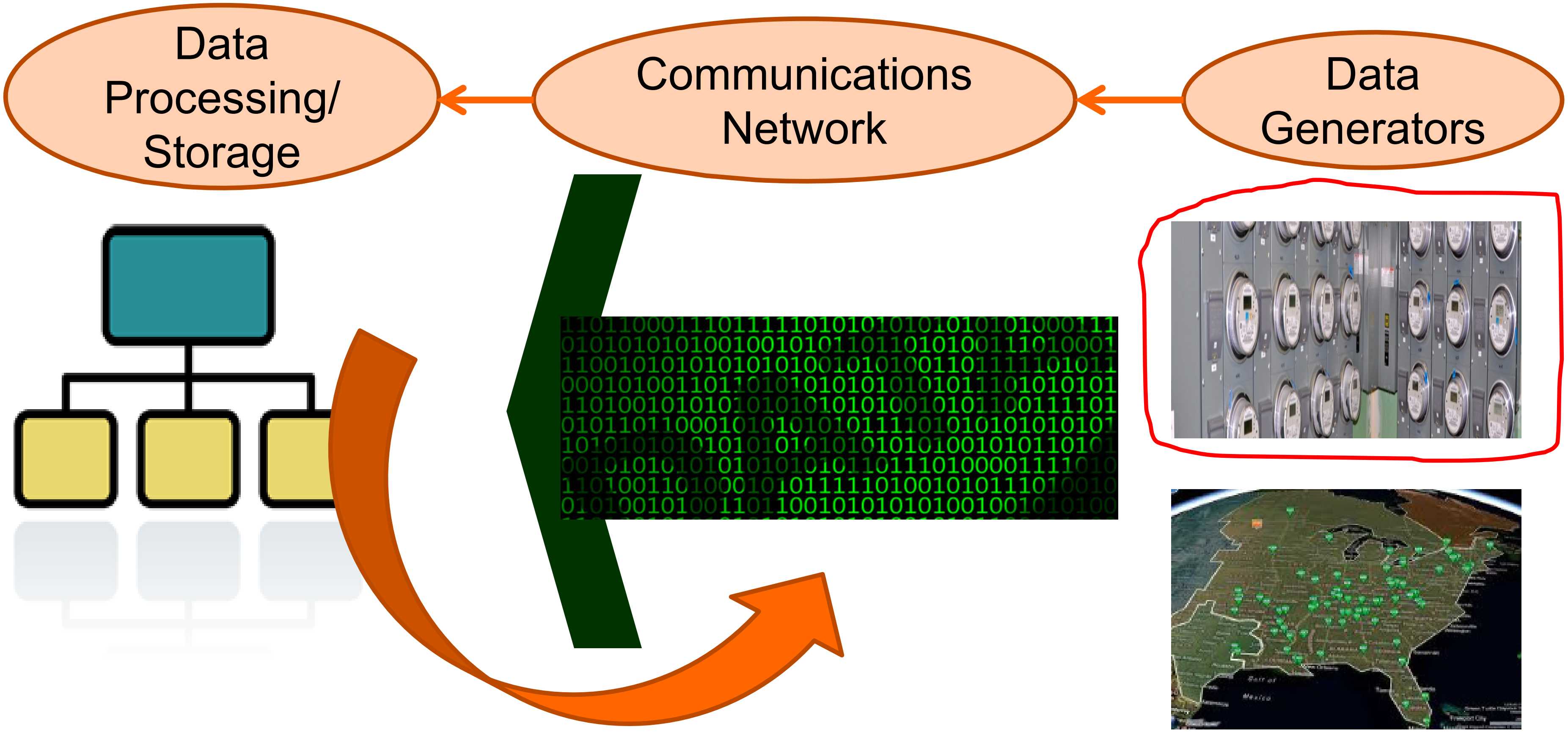
- Network capacity is a precious resource for electric utilities because they are
  - leasing networks from third-party providers or
  - building infrastructure themselves and leasing bandwidth out

- W. Luan, D. Sharp, and S. Lancashire, “Smart grid communication network capacity planning for power utilities”
- Engage Consulting Limited, “High-level smart meter data traffic analysis,” for the Energy Networks Association(ENA).
- “Arcadian’s Smart Grid: Licensed Spectrum Network to Own or Rent,” Greentechgrid.
- M. Kennedy, “Leveraging investment in fiber optic communications,” IEEE Smart Grid.

Calls for a Data Volume Concentration Process



Data Stream – An Abstract View



Need In-Network Data Processing/Analytics

Some Solution Approaches

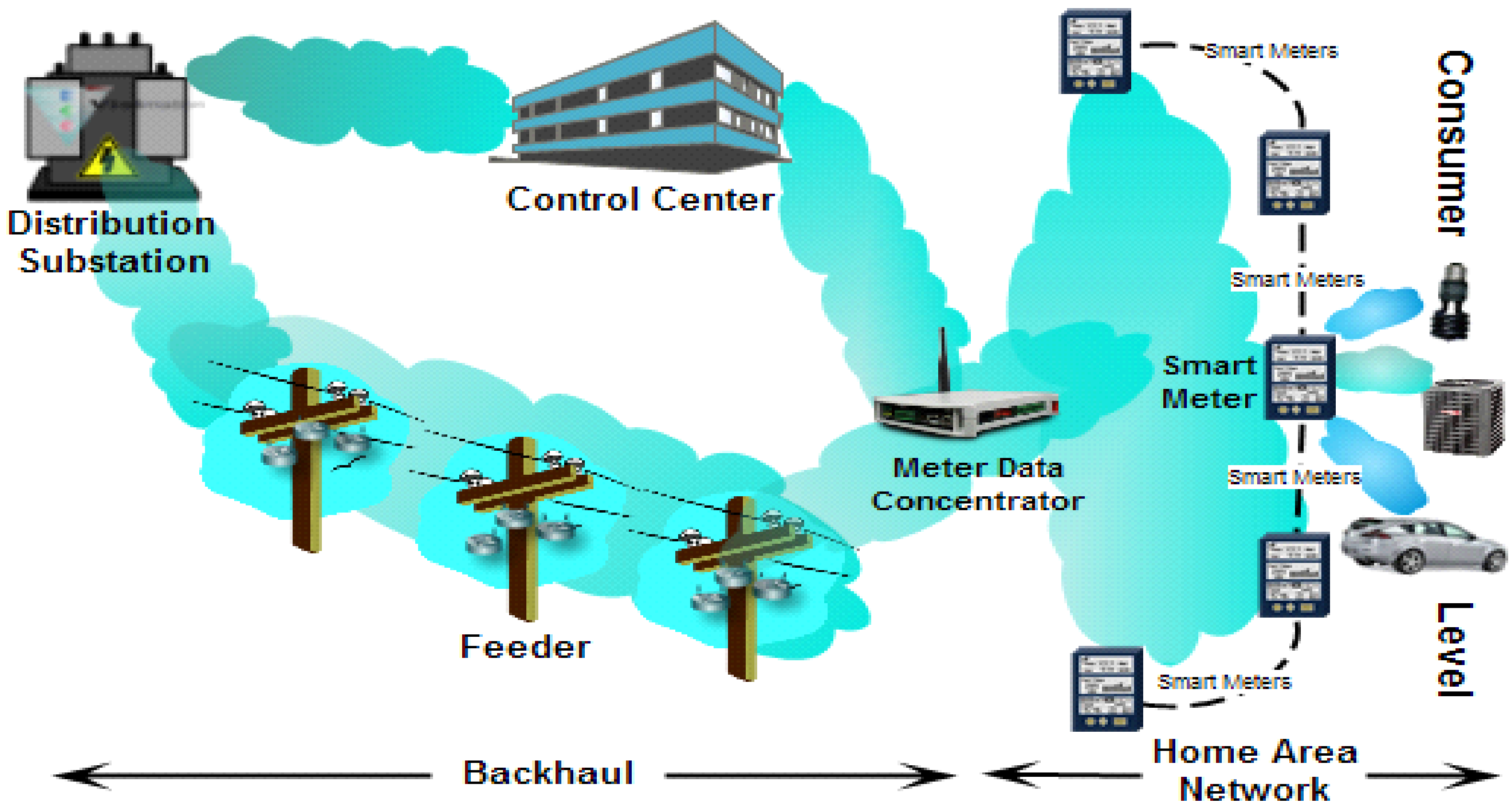
- Message Concatenation: group multiple messages together with a common protocol header as opposed to each message having its own header
- Aggregation: apply aggregation functions as appropriate based on requirements to combine data from several measurement
- Compress: exploit statistical redundancy to represent data more concisely without losing information

Role of Data Concentrator Units

- Approaches listed have the following two features in common.
  - They must leverage the information needs of grid operators for minimizing information volume while meeting any QoS constraints at the same time.
  - They require mechanisms at intermediate relay points to execute algorithms to meet their information volume concentration objectives.

vs

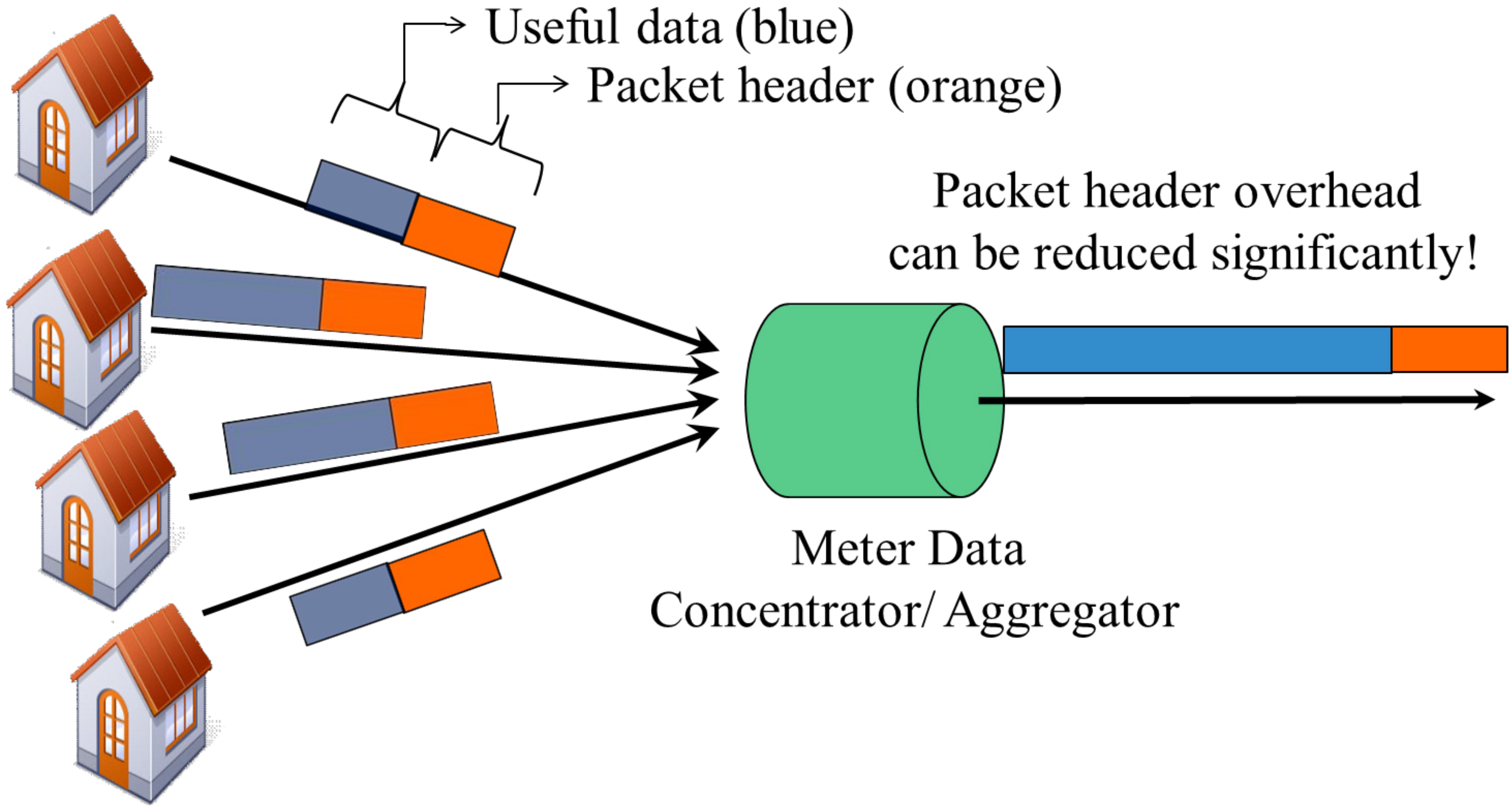
Power Distribution System



Position in Existing Literature

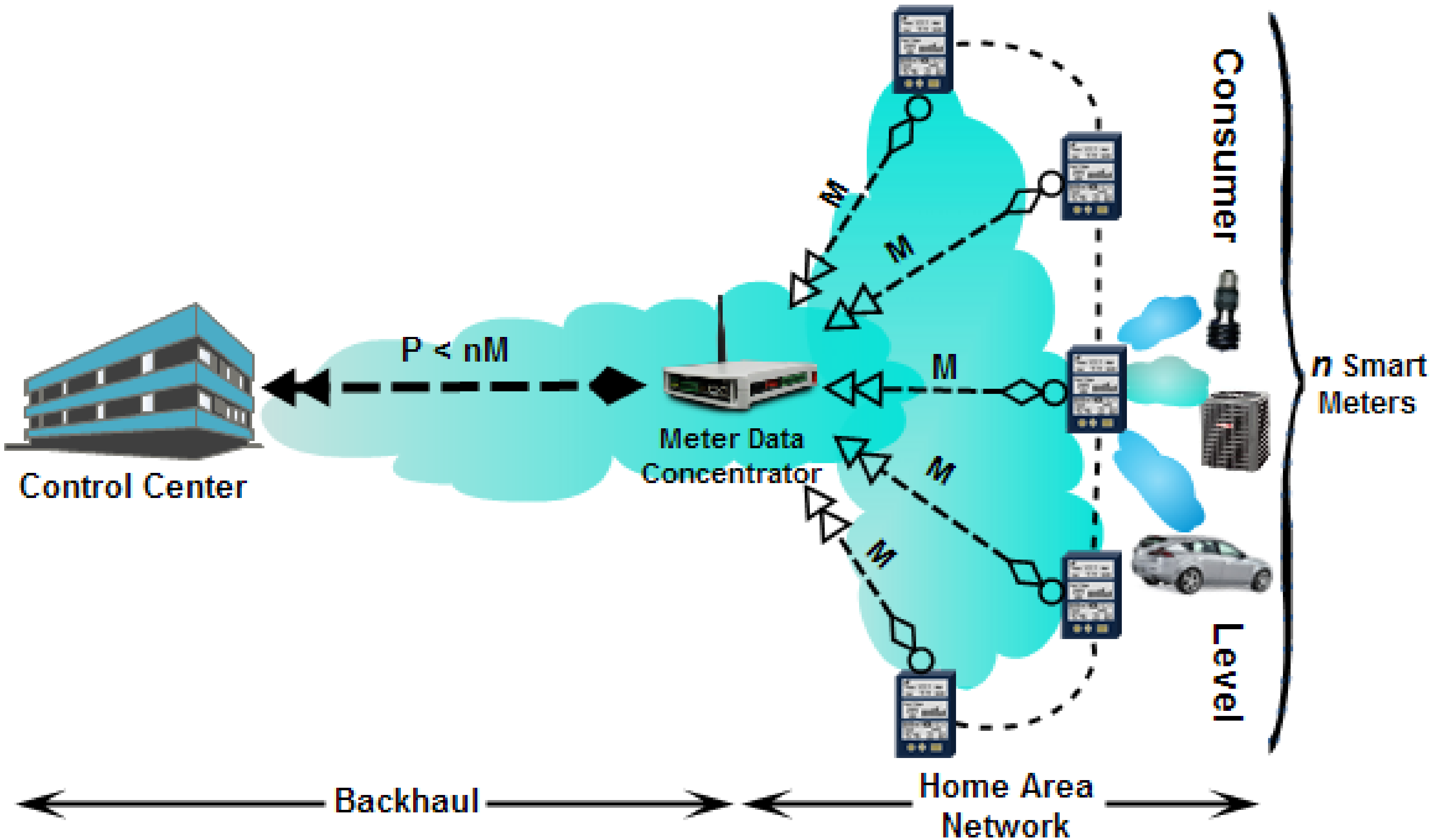
- Data aggregation in the field of WSNs
  - Access to electric power at all times with backup batteries
  - Shifts the focus of the problem from battery life of nodes to reduction of network capacity utilization
- Data concentration for WAMPAC
  - Similar work missing related to power distribution system and metering

Smart Metering Message Concatenation (SMMC)



Karimi, B., Namboodiri, V., Jadliwala, M., "Scalable Meter Data Collection in Smart Grids through Message Concatenation," *IEEE Transactions on Smart Grids*, vol. 6, pp. 1697–1706, 2015.

SMMC Problem and Desired Result



Smart Meter Data Messages Types

Message/ Traffic Description	Size (Bytes)	Delay Obj.
Meter clock synchronization	64	2 secs
Interval data read	480	Best effort
Firmware patch/upgrade confirmation/ acknowledge	20	Best effort
Meter ping (on demand read)	64	2 secs
Meter remote diagnostic	500	2 secs
Tamper notification	64	5 secs
Meter remote disconnect/ reconnect response	500	2 secs

W.Luan, D.Sharp, and S.Lancashire,"Smart grid communication network capacity planning for power utilities," in Transmission and Distribution Conference and Exposition, 2010 IEEE PES, April 2010.

Smart Meter Data Message

- Meter id, Equipment status, Type of message, metered data (ANSI C12 standard)
- The size of the packet header depends on the communication protocol used.
  - It can add from 10 to 50 bytes per packet of data sent.

Transport Protocols

Protocols:	Estimated Protocol frame size (bytes)	Used For		Physical Carriers						
		Remote AMR	AMI	PSTN	GSM	GPRS	Low Power RF	PLC Broadband[6]	PLC Narrowband	Mesh radio
IEC 61334 (S-FSK, FSK, OFDM)	45 bytes[7]	Y	Y		-	-	-	Y	Y[8]	-
IEC 62056-31 Euridis[9]	45 bytes[10]	Y	-		-	-	-	-	-	-
EN 13757 M-Bus[11]	27 bytes	-	Y[12]		-	-	Y	-	-	-
TCP/IP protocol[13]	50 bytes	Y	Y	Y	Y	Y		Y	Y	
SITRED[14]	45 bytes[15]	Y	Y					Y	Y	
PRIME	8 bytes[16]	Y	Y						Y	
EverBlu	unknown[17]	Y	Y			Y				Y
OPERA/UPA[18]	24 bytes[19]	Y	Y	-	-	-		Y	Y	-
ZigBee[20]	15 bytes		Y				Y			Y

Heuristic Approach

- Earliest Deadline First (EDF)

- DCU queues all arriving packets
- When time = time of deadline for message with earliest deadline
  - Insert message into newly created packet
  - Fill it with other messages to maximize packet size using one of the schemes presented next

- Deadline at DCU can be created by subtracting off some estimate of network and processing latencies from end-point deadline.

Some Possible Heuristics

Algorithm	Description
EDF-DKB	Inserts deadline messages as much as possible inside the packet and the remaining space will be filled through knapsack selection over best-effort messages that have been queued.
EDF-SDKB	Only a single deadline message sits inside the packet with any available space filled with non-deadline messages in the non-deadline queue through knapsack selection.
EDF-FCFS	Messages will be placed in the packet according to their arrival sequence from a common queue of deadline and non-deadline messages on a first-come first-served basis.
EDF-KN	Messages are chosen from a common pool of deadline and best-effort messages selected through the knapsack algorithm.
EDF-KDKB	A sequence of knapsack selections first on all queued deadline messages and then over the queued best-effort messages if needed to fill the packet.
EDF-KBKD	Reverse order of knapsack process in EDF-KDKB working first on the queued best-effort messages and then on the deadline messages if needed.

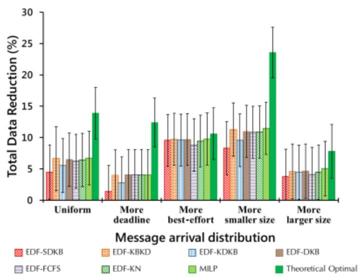
Evaluation Methodology

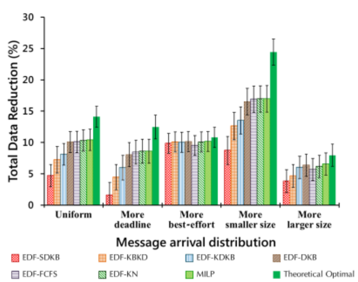
- Reference Algorithms
  - EDF-based MILP
  - Theoretical
- Discrete Event Simulations in MATLAB
  - 95% confidence intervals shown
- Message generation at meters assumed to follow a Poisson process with an aggregate arrival rate of λ messages/sec at the DCU
  - λ of 0.1, 0.5 and 1 studied
- Mix of message types arriving at DCU
  - Assume a Beta distribution with shape parameters $\alpha > 0$, and $\beta > 0$ that can be varied to vary the mix
- For simplicity we have assumed there is no *Network* and *Processing delay* in these evaluations.

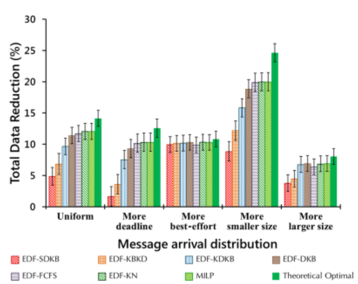
Message Type Distributions

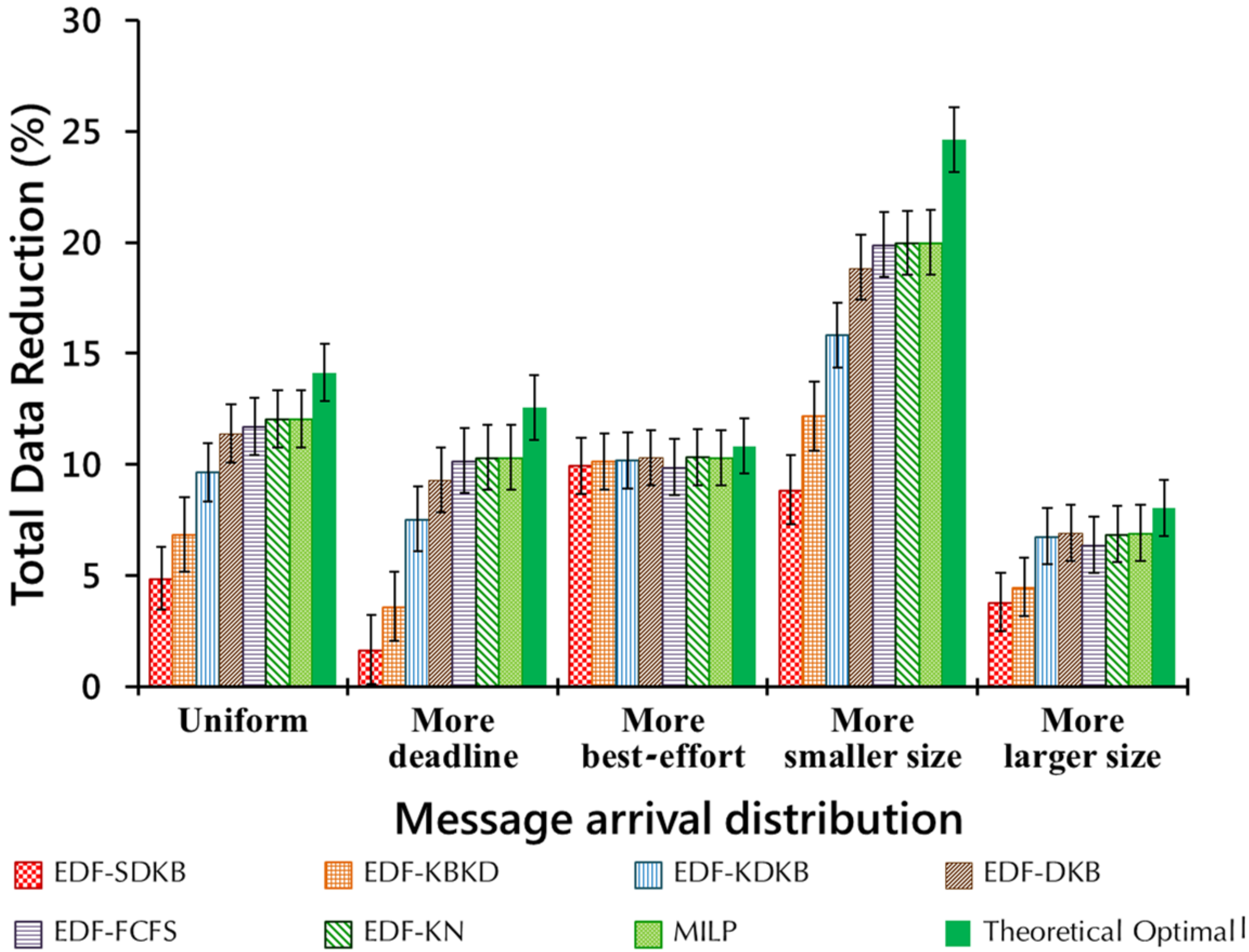
Distribution	Description
Uniform ($\alpha = 1$, $\beta = 1$)	The traffic would have almost equal percentage of all message types.
More smaller ($\alpha = 2.8$, $\beta = 1.9$)	Most of the arrived messages are of the smaller size of message types.
More larger ($\alpha = 0.18$, $\beta = 0.25$)	There is higher percentage of large message size and very few numbers of small size messages.
More deadline ($\alpha = 1$, $\beta = 1.8$)	Most of the times there are incoming messages with deadline restriction.
More best-effort ($\alpha = 2.5$, $\beta = 0.5$)	There are very few numbers of messages with a deadline and so many best-effort messages.

Data reduction ratio vs Different message arrival rates (λ=0.1, 0.5, and 1)









Application-Aware Data Aggregation

- Data from multiple samples could be represented by one sample

- Aggregation function examples {2,5,8,5}
  - Minimum: 2
  - Maximum: 8
  - Average: 4
  - Most Repeated: 5
  - Median: 5

V. C. Dev, U. Das, V. Namboodiri, S. Chakraborty, V. Aravinthan, Y. Guo, and A. Srivastava, "Towards application-aware data concentration schemes for advanced metering infrastructures," to appear, 2015 *IEEE SmartGridComm*, November 2015.

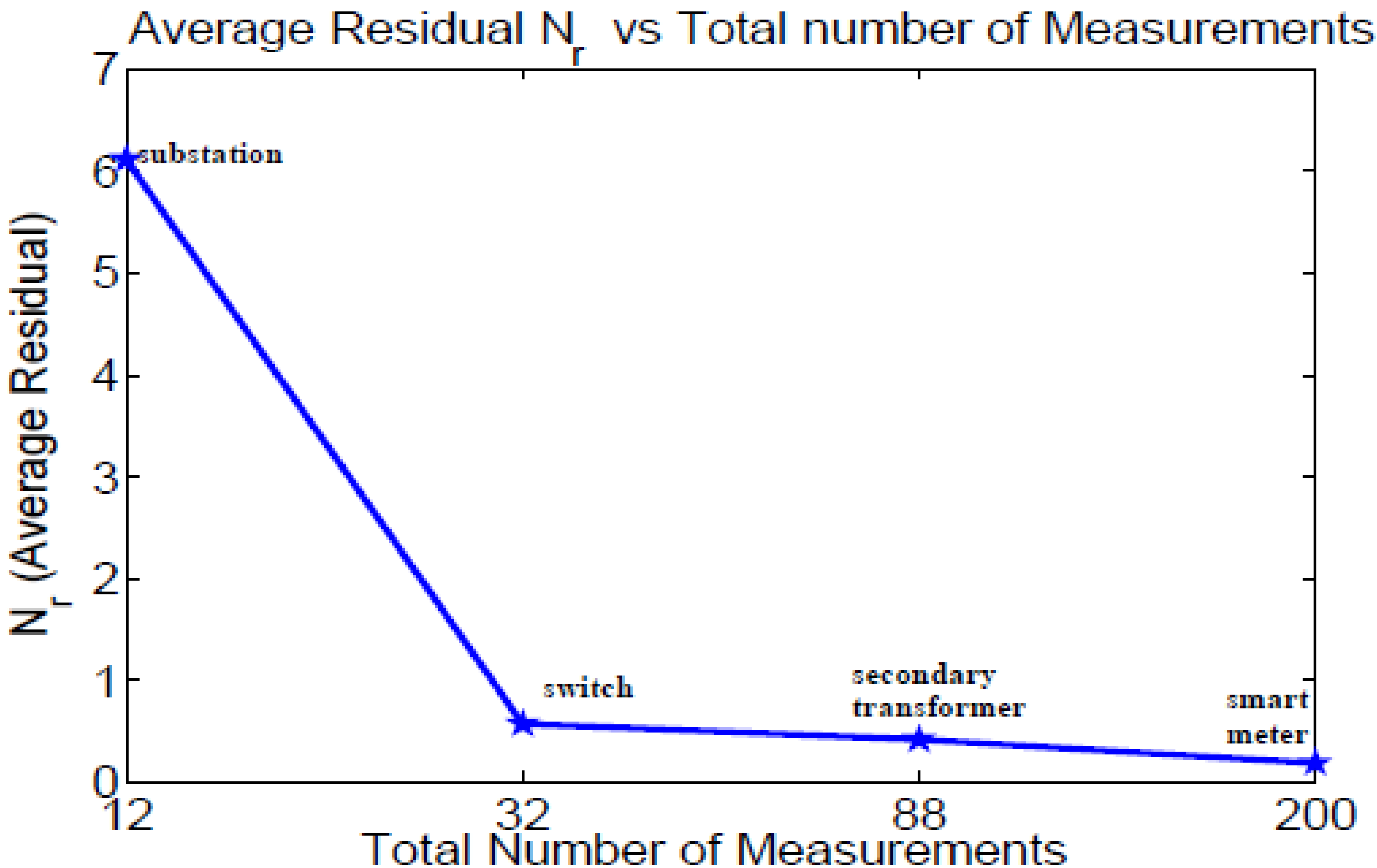
Data Granularity

- A measure of how well the original data samples are represented after aggregation

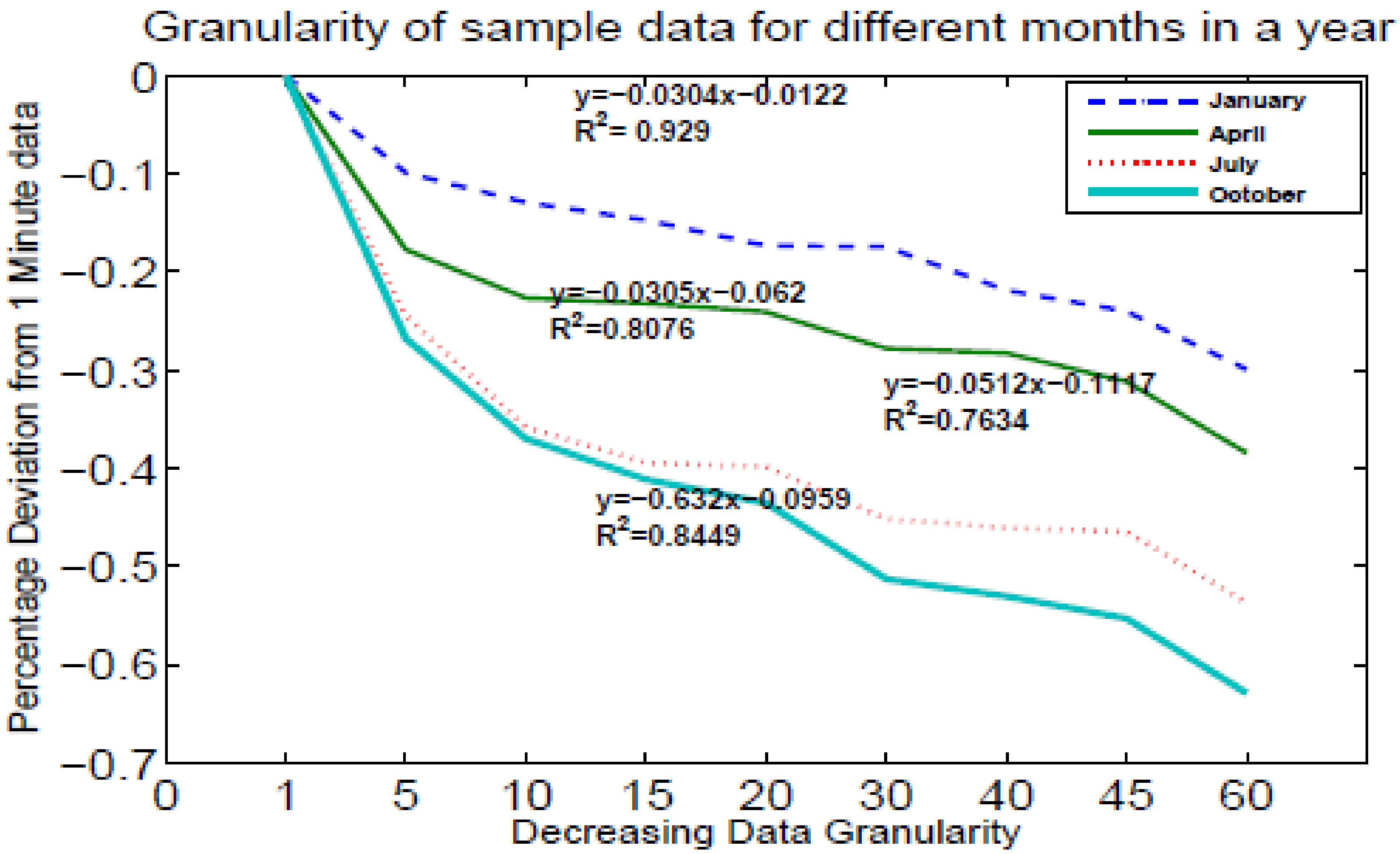
Assume that each smart meter s_i generates m data samples each of size g_{ij} bytes each over a period of time T and sends it to the DCU. Based on network congestion, assume that the data concentrator aggregates multiples of incoming messages from all smart meter into a total of k messages $a_{i1} \cdots a_{ik}$, $1 \leq k \leq n \cdot m$ before sending them out over the backhaul network.[1] We can then define data granularity γ over period T as

$$\gamma = 1 - \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} g_{ij} - \sum_{i=1}^{k} a_i}{\sum_{i=1}^{n} \sum_{j=1}^{m} g_{ij}} \qquad (1)$$

Application: State Estimation



Application: Transformer Life Estimation

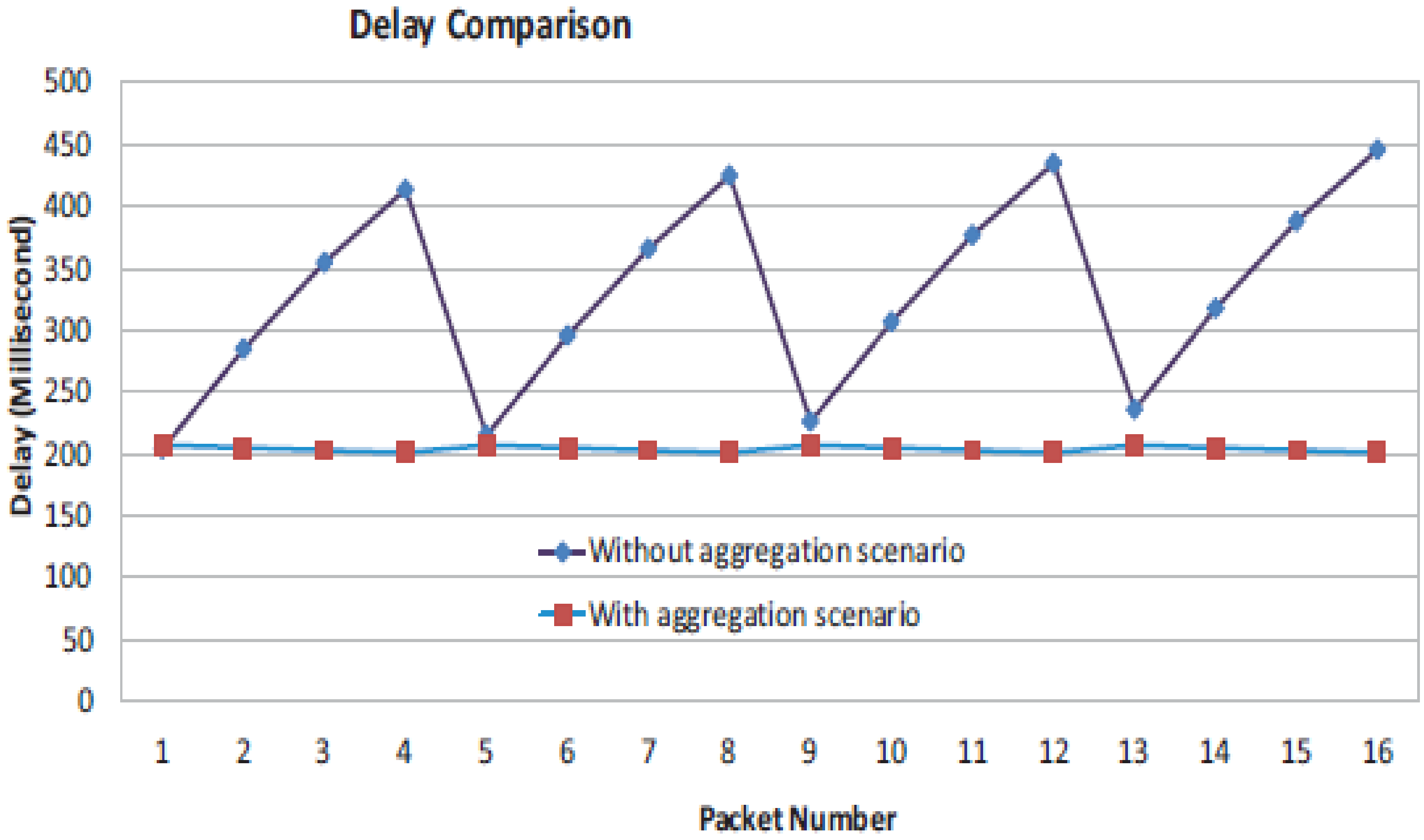


Algorithm

Algorithm 1 Aggregation Algorithm at DCU

Take algorithm execution period T, initial estimate of backhaul network capacity C, application data granularity tolerance level γ_{lb} as inputs
Repeat every T time units
Compute τ using knowledge of all samples that arrived over the last period and a current estimate of C
if $\tau > 1$ **then**
    Collect a window of τ samples, apply aggregation function, then send out
    Compute data granularity γ using knowledge of all incoming samples and outgoing aggregate messages
    if $\gamma < \gamma_{lb}$ **then**
        Notify control center of compromise in data quality below specified threshold
    end if
end if

Preliminary Results



Recap and Conclusions

- Efforts to manage data volume needed in conjunction with design of communications infrastructures and new applications
- Just message concatenation can reduce data volume of metering applications by 10-25%
- Data aggregation is another effective mechanism if the application-quality tradeoff can be managed well

Future Work

- Distributed data management and analytics, beginning with:
  - What data needs to be collected
  - What granularity data is needed where
- Data concentration and aggregation of PMU data
- Employ compression techniques in addition to aggregation and concatenation

Thank you!