Randomized Time Division strategy for broadcasting information

Chandra Nair

Department of Information Engineering

The Chinese University of Hong Kong
chandra.nair@gmail.com

Feb 26, 2014

BROADCAST CHANNELS [COVER '72]

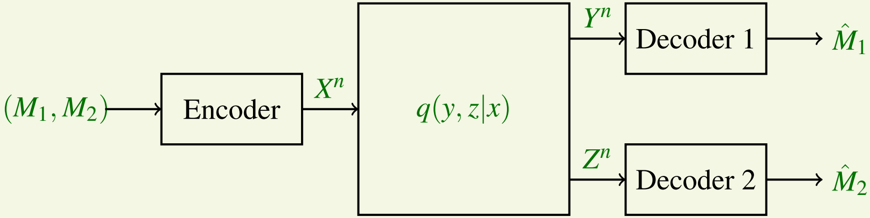

Figure : Discrete memoryless broadcast channel

- How to simultaneously communicate independent messages to the two receivers over a shared medium?
- Want to take advantage of the fact that the transmitted symbols are corrupted differently for the two receivers

Encoding strategy [Cover '72]

A natural coding strategy is TIME DIVISION.

Can one do better?

Imagine that you are simultaneously speaking to two people:

- Johnny English understands only English alphabets: A, B and when it is not either he denotes it as $*$.
- Ms. Discrete understands understands only 0 and 1, and when it is neither she denotes it as ?.

Suppose you want to communicate a word AABBAABB to Johnny English and the word 101110 to Ms. Discrete then your can either speak

- AABBAABB101110 or 101110AABBAABB

ENCODING STRATEGY [COVER '72]

A natural coding strategy is TIME DIVISION.

Can one do better?

Imagine that you are simultaneously speaking to two people:

- Johnny English understands only English alphabets: A, B and when it is not either he denotes it as $*$.
- Ms. Discrete understands understands only 0 and 1, and when it is neither she denotes it as ?.

Suppose you want to communicate a word AABBAABB to Johnny English and the word 101110 to Ms. Discrete then your can either speak

- AABBAABB101110 or 101110AABBAABB

Encoding strategy [Cover '72]

A natural coding strategy is TIME DIVISION.

Can one do better?

Imagine that you are simultaneously speaking to two people:

- Johnny English understands only English alphabets: A, B and when it is not either he denotes it as $*$.
- Ms. Discrete understands understands only 0 and 1, and when it is neither she denotes it as $?$.

Suppose you want to communicate a word AABBAABB to Johnny English and the word 101110 to Ms. Discrete then your can either speak

- AABBAABB101110 or 101110AABBAABB
- In particular you can choose any one of the $\binom{14}{8} = 3003$ possibilities

Which particular order you choose can convey $\lfloor \log_2(3003) \rfloor = 11$ additional bits of information that is decodable by both receivers.
Such a strategy is called RANDOMIZED TIME DIVISION

Comparing achievable rate regions

Using time-division one can achieve

$$R_1 \leq \alpha C_{JE} = \alpha \log_2(3)$$
$$T_2 \leq (1-\alpha) C_D = (1-\alpha)\log_2(3)$$

for any $0 \leq \alpha \leq 1$.

Using randomized time-division, one can achieve (R_1, R_2) satisfying

$$R_1 \leq H(\alpha) + \alpha \log_2(2)$$
$$R_2 \leq H(\alpha) + (1-\alpha)\log_2(2)$$
$$R_1 + R_2 \leq H(\alpha) + \alpha \log_2(2) + (1-\alpha)\log_2(2)$$

for any $0 \leq \alpha \leq 1$.

Rate Regions

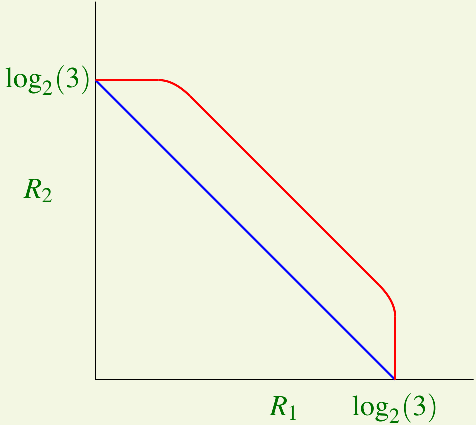

Figure : Achievable regions for time division vs randomized time division

Recap: Randomized time-division strategy

- One picks a codebook of binary codewords that is decodable by both receivers
- The codeword is picked from this codebook that reveals *common (timing) information* to both receivers
- On the locations where this *timing* codeword has 0, the encoder sends a codeword tailored to be decoded by receiver Y
- On the locations where this *timing* codeword has 1, the encoder sends a codeword tailored to be decoded by receiver Z

Using randomized time-division, one can achieve (R_1, R_2) satisfying

$$R_1 \leq I(W;Y) + \mathrm{P}(W=0)I(X;Y|W=0)$$
$$R_2 \leq I(W;Z) + \mathrm{P}(W=1)I(X;Z|W=1)$$
$$R_1 + R_2 \leq \min\{I(W;Y), I(W;Z)\} + \mathrm{P}(W=0)I(X;Y|W=0) + \mathrm{P}(W=1)I(X;Z|W=1)$$

Recap: Randomized time-division strategy

- One picks a codebook of binary codewords that is decodable by both receivers
- The codeword is picked from this codebook that reveals *common (timing) information* to both receivers
- On the locations where this *timing* codeword has 0, the encoder sends a codeword tailored to be decoded by receiver Y
- On the locations where this *timing* codeword has 1, the encoder sends a codeword tailored to be decoded by receiver Z

Using randomized time-division, one can achieve (R_1, R_2) satisfying

$$
\begin{aligned}
R_1 &\leq I(W;Y) + \mathrm{P}(W=0)I(X;Y|W=0) \\
R_2 &\leq I(W;Z) + \mathrm{P}(W=1)I(X;Z|W=1) \\
R_1 + R_2 &\leq \min\{I(W;Y), I(W;Z)\} + \mathrm{P}(W=0)I(X;Y|W=0) \\
&\quad + \mathrm{P}(W=1)I(X;Z|W=1)
\end{aligned}
$$

How good is this strategy?

Randomized time divison is a naive improvement over time-division.

The best achievable region for broadcast channel is due to Marton [Mar '79]

Marton's achievable region

The union of rate pairs (R_1, R_2) such that the rate pairs satisfy

$$\begin{aligned} R_1 &\leq I(UW;Y) \\ R_2 &\leq I(VW;Z) \\ R_1 + R_2 &\leq \min\{I(W;Y), I(W;Z)\} + I(U;Y|W) \\ &\quad + I(V;Z|W) - I(U;V|W) \end{aligned}$$

for any $(U,V,W) - X - (Y,Z)$ is achievable.

Issue: The above region was not computable (no cardinality bounds on U,V,W)

Investigating Marton's achievable region

- Conjectured that for a particular binary input broadcast channel

$$SR_{Marton} = SR_{RTD} \qquad \text{[Nair-Wang '08]}$$

- Established cardinality bounds

$$|U| \leq |X|, |V| \leq |X|, |W| \leq |X| + 4$$

for evaluating Marton's achievable region [Gohari-Anantharam '09]
    - Novel kind of perturbation arguments
- Extended the perturbation arguments to show that the earlier conjectured result is true [Jog-Nair '09]
- Showed that for all *binary* input broadcast channels

$$SR_{Marton} = SR_{RTD} \qquad \text{[Geng-Nair-Wang '10]}$$

INVESTIGATING MARTON'S ACHIEVABLE REGION

- Conjectured that for a particular binary input broadcast channel

$$SR_{Marton} = SR_{RTD} \qquad \text{[Nair-Wang '08]}$$

- Established cardinality bounds

$$|U| \leq |X|, |V| \leq |X|, |W| \leq |X| + 4$$

  for evaluating Marton's achievable region [Gohari-Anantharam '09]
  - Novel kind of perturbation arguments
- Extended the perturbation arguments to show that the earlier conjectured result is true [Jog-Nair '09]
- Showed that for all *binary* input broadcast channels

$$SR_{Marton} = SR_{RTD} \qquad \text{[Geng-Nair-Wang '10]}$$

Investigating Marton's achievable region

- Conjectured that for a particular binary input broadcast channel

$$SR_{Marton} = SR_{RTD} \qquad \text{[Nair-Wang '08]}$$

- Established cardinality bounds

$$|U| \leq |X|, |V| \leq |X|, |W| \leq |X| + 4$$

  for evaluating Marton's achievable region [Gohari-Anantharam '09]
  - Novel kind of perturbation arguments
- Extended the perturbation arguments to show that the earlier conjectured result is true [Jog-Nair '09]
- Showed that for all *binary* input broadcast channels

$$SR_{Marton} = SR_{RTD} \qquad \text{[Geng-Nair-Wang '10]}$$

Investigating Marton's achievable region

- Conjectured that for a particular binary input broadcast channel

$$SR_{Marton} = SR_{RTD}$$ [Nair-Wang '08]

- Established cardinality bounds

$$|U| \leq |X|, |V| \leq |X|, |W| \leq |X| + 4$$

  for evaluating Marton's achievable region [Gohari-Anantharam '09]
  - Novel kind of perturbation arguments
- Extended the perturbation arguments to show that the earlier conjectured result is true [Jog-Nair '09]
- Showed that for all *binary* input broadcast channels

$$SR_{Marton} = SR_{RTD}$$ [Geng-Nair-Wang '10]

Investigating Marton's achievable region

- Conjectured that for a particular binary input broadcast channel

$$SR_{Marton} = SR_{RTD} \qquad \text{[Nair-Wang '08]}$$

- Established cardinality bounds

$$|U| \leq |X|, |V| \leq |X|, |W| \leq |X| + 4$$

  for evaluating Marton's achievable region [Gohari-Anantharam '09]
    - Novel kind of perturbation arguments
- Extended the perturbation arguments to show that the earlier conjectured result is true [Jog-Nair '09]
- Showed that for all *binary* input broadcast channels

$$SR_{Marton} = SR_{RTD} \qquad \text{[Geng-Nair-Wang '10]}$$

What about entire region and not just sum-rate?

IMPROVED CARDINALITY BOUNDS

Combining

- perturbation arguments [Gohari-Anantharam]
- concave envelope interpretation for extremal auxiliaries [Nair]

one can show that it suffices to consider auxiliairies U, V that satisfy

$$|U| + |V| \leq |X| + 1$$

to evaluate Marton's achievable region [Anantharam-Gohari-Nair '12]

This shows that when $|X| = 2$, $|U| + |V| \leq 3$.
Since $|U|, |V| \geq 1$; the only possibilities are $|U| = 2, |V| = 1$ or $|V| = 2, |U| = 1$.
This in particular implies that

- Marton's achievable region *matches* randomized time division when $|X|$ is binary

Improved cardinality bounds

Combining

- perturbation arguments [Gohari-Anantharam]
- concave envelope interpretation for extremal auxiliaries [Nair]

one can show that it suffices to consider auxiliairies U, V that satisfy

$$|U| + |V| \leq |X| + 1$$

to evaluate Marton's achievable region [Anantharam-Gohari-Nair '12]

This shows that when $|X| = 2$, $|U| + |V| \leq 3$.
Since $|U|, |V| \geq 1$; the only possibilities are $|U| = 2, |V| = 1$ or $|V| = 2, |U| = 1$.
This in particular implies that

- Marton's achievable region *matches* randomized time division when $|X|$ is binary

MAIN RESULT

$$T_{q,\alpha}(X) := \sup_{(U,V)\to X\to(Y,Z)} \alpha I(U;Y) + I(V;Z) - I(U;V)$$

Cardinality bounds on U, V

For all broadcast channels $q(y,z|x)$, for all $\lambda \in [0,1]$, for all $\alpha \geq 1$, to compute

$$\mathsf{C}[-(\alpha-\lambda)H(Y) - \lambda H(Z) + T_{q,\alpha}(X)]$$

it suffices to consider $\|U\| + \|V\| \leq \|X\| + 1$.

Corollary

1. Suffices to consider $\|U\| + \|V\| \leq \|X\| + 1$, $\|W\| \leq \|X\| + 4$ to compute $\mathrm{MIB}(q)$

Main result

$$T_{q,\alpha}(X) := \sup_{(U,V)\to X\to (Y,Z)} \alpha I(U;Y) + I(V;Z) - I(U;V)$$

Cardinality bounds on U, V

For all broadcast channels $q(y,z|x)$, for all $\lambda \in [0,1]$, for all $\alpha \geq 1$, to compute

$$\mathsf{C}[-(\alpha-\lambda)H(Y) - \lambda H(Z) + T_{q,\alpha}(X)]$$

it suffices to consider $\|U\| + \|V\| \leq \|X\| + 1$.

Corollary

1. Suffices to consider $\|U\| + \|V\| \leq \|X\| + 1$, $\|W\| \leq \|X\| + 4$ to compute $\text{MIB}(q)$

IDEA OF PROOF

Suppose $p(u,v,x)$ is an *extremal distribution* such that

$$\mathsf{C}[-(\alpha-\lambda)H(Y)-\lambda H(Z)+T_{q,\alpha}(X)]$$
$$=-(\alpha-\lambda)H(Y)-\lambda H(Z)+\alpha I(U;Y)+I(V;Z)-I(U;V),$$

then the r.h.s. is *locally concave* with respect to all perturbations of $p(u,v,x)$. Rearrange the right hand side as

$$\lambda(H(Y)-H(Z))-\alpha H(Y|U)+H(V|U)-H(Z|V)$$

Consider a perturbation of the form

$$p_\epsilon(u,v,x)=p(u,v,x)(1+\epsilon f(u)), \qquad \Big(\sum_u p(u)f(u)=0\Big).$$

For the second derivative to be negative, it must be that the second derivative of the term

$$\frac{d^2}{d\epsilon^2}\left[H(Y)-H(Z)\right]_{\epsilon=0}\le 0$$

IDEA OF PROOF

Suppose $p(u,v,x)$ is an *extremal distribution* such that

$$\begin{aligned} &\mathsf{C}[-(\alpha-\lambda)H(Y)-\lambda H(Z)+T_{q,\alpha}(X)] \\ &\quad = -(\alpha-\lambda)H(Y)-\lambda H(Z)+\alpha I(U;Y)+I(V;Z)-I(U;V), \end{aligned}$$

then the r.h.s. is *locally concave* with respect to all perturbations of $p(u,v,x)$. Rearrange the right hand side as

$$\lambda(H(Y)-H(Z))-\alpha H(Y|U)+H(V|U)-H(Z|V)$$

Consider a perturbation of the form

$$p_\epsilon(u,v,x)=p(u,v,x)(1+\epsilon f(u)), \qquad \Big(\sum_u p(u)f(u)=0\Big).$$

For the second derivative to be negative, it must be that the second derivative of the term

$$\frac{d^2}{d\epsilon^2}\left[H(Y)-H(Z)\right]_{\epsilon=0}\le 0$$

Idea of proof (cntd...)

Alternately, rearrange the right hand side as

$$(1-\lambda)(H(Z)-H(Y)) - H(Z|V) + H(U|V) - H(U|Y) - (\alpha-1)H(Y|U)$$

Consider a perturbation of the form

$$\hat{p}_\epsilon(u,v,x) = p(u,v,x)(1+\epsilon g(v)), \qquad \Big(\sum_v p(v)g(v) = 0\Big).$$

For the second derivative to be negative, it must be that the second derivative of the term

$$\frac{d^2}{d\epsilon^2}\left[H(Z)-H(Y)\right]_{\epsilon=0} \leq 0$$

Observation

For a fixed channel $q(y,z|x)$ the term $H(Y) - H(Z)$ depends only on $p(x)$.

Hence, if there exists $f(u)$ and $g(v)$ such that $p_\epsilon(x) = \hat{p}_\epsilon(x)$ for all $x \in \mathcal{X}$, then one would need to have

$$\frac{d^2}{d\epsilon^2}\left[H(Y) - H(Z)\right]_{\epsilon=0} = 0.$$

This will in turn force the convex terms to have zero second derivative as well.

As a consequence, it will turn out that the expression

$$-(\alpha - \lambda)H(Y) - \lambda H(Z) + \alpha I(U;Y) + I(V;Z) - I(U;V)$$

will remain unchanged by either of these perturbations.

Set ϵ large enough so that the support of U or V reduces by one.

OBSERVATION

For a fixed channel $q(y,z|x)$ the term $H(Y) - H(Z)$ depends only on $p(x)$.

Hence, if there exists $f(u)$ and $g(v)$ such that $p_\epsilon(x) = \hat{p}_\epsilon(x)$ for all $x \in \mathcal{X}$, then one would need to have

$$\frac{d^2}{d\epsilon^2}\left[H(Y) - H(Z)\right]_{\epsilon=0} = 0.$$

This will in turn force the convex terms to have zero second derivative as well.

As a consequence, it will turn out that the expression

$$-(\alpha - \lambda)H(Y) - \lambda H(Z) + \alpha I(U;Y) + I(V;Z) - I(U;V)$$

will remain unchanged by either of these perturbations.

Set ϵ large enough so that the support of U or V reduces by one.

Observation

For a fixed channel $q(y,z|x)$ the term $H(Y)-H(Z)$ depends only on $p(x)$.

Hence, if there exists $f(u)$ and $g(v)$ such that $p_\epsilon(x) = \hat{p}_\epsilon(x)$ for all $x \in \mathcal{X}$, then one would need to have

$$\frac{d^2}{d\epsilon^2}\left[H(Y)-H(Z)\right]_{\epsilon=0} = 0.$$

This will in turn force the convex terms to have zero second derivative as well.

As a consequence, it will turn out that the expression

$$-(\alpha-\lambda)H(Y) - \lambda H(Z) + \alpha I(U;Y) + I(V;Z) - I(U;V)$$

will remain unchanged by either of these perturbations.

Set ϵ large enough so that the support of U or V reduces by one.

Conditions for existence of $f(u), g(v)$

1. $\sum_{u,v} p(u,v,x) f(u) = \sum_{u,v} p(u,v,x) g(v) \ \ \forall x \in \mathcal{X}$.
   - From the condition: $p_\epsilon(x) = \hat{p}_\epsilon(x)$ for all $x \in \mathcal{X}$.
2. $\sum_{u,v,x} p(u,v,x) f(u) = 0$.
   - From the condition: $p_\epsilon(x)$ is a valid probability distribution.
3. $\sum_{u,v,x} p(u,v,x) g(v) = 0$.
   - From the condition: $\hat{p}_\epsilon(x)$ is a valid probability distribution.

So there are $\|X\| + 1$ linear constraints on a vector of size $\|U\| + \|V\|$.

A non-trivial solution exists when $\|U\| + \|V\| > \|X\| + 1$.

Comments

It is reasonably straightforward to map this bound on auxiliaries to Marton's inner bound.

The bound $\|W\| \leq \|X\| + 4$ comes from convexification (traditional Caratheodory based arguments)

This bound seems to indicate a trade-off between communicating with one receiver vs. the other

- In some sense, the essence of broadcast channel

One can extract more properties of *extremal* $p(u,v,x)$ from the argument: consider the Hessian H (of size $\|X\| - 1$) of the mapping $p(x) \mapsto H(Y) - H(Z)$

- $\|U\| - 1$ is at most number of negative eigenvalues of H
- $\|V\| - 1$ is at most number of positive eigenvalues of H

COMMENTS

It is reasonably straightforward to map this bound on auxiliaries to Marton's inner bound.

The bound $\|W\| \leq \|X\| + 4$ comes from convexification (traditional Caratheodory based arguments)

This bound seems to indicate a trade-off between communicating with one receiver vs. the other

- In some sense, the essence of broadcast channel

One can extract more properties of *extremal* $p(u, v, x)$ from the argument: consider the Hessian H (of size $\|X\| - 1$) of the mapping $p(x) \mapsto H(Y) - H(Z)$

- $\|U\| - 1$ is at most number of negative eigenvalues of H
- $\|V\| - 1$ is at most number of positive eigenvalues of H

Conclusion

- Cover introduced the broadcast channel
- In that he talked about the Spanish-English channel (randomized time-division)

- Later Marton and Cover-van der Meulen came up with improved achievable regions using auxiliary random variables

- After trying hard and determining the extremal auxiliary random variables, we see that regions revert to randomized time-division (binary) and an as yet undetermined generalization of randomized time-division.

- May be there is a story behind extremal auxiliary random variables that will shed light into decades long open problems.

Thank You

CONCLUSION

- Cover introduced the broadcast channel
- In that he talked about the Spanish-English channel (randomized time-division)
- Later Marton and Cover-van der Meulen came up with improved achievable regions using auxiliary random variables
- After trying hard and determining the extremal auxiliary random variables, we see that regions revert to randomized time-division (binary) and an as yet undetermined generalization of randomized time-division.
- May be there is a story behind extremal auxiliary random variables that will shed light into decades long open problems.

Thank You

Conclusion

- Cover introduced the broadcast channel
- In that he talked about the Spanish-English channel (randomized time-division)

- Later Marton and Cover-van der Meulen came up with improved achievable regions using auxiliary random variables

- After trying hard and determining the extremal auxiliary random variables, we see that regions revert to randomized time-division (binary) and an as yet undetermined generalization of randomized time-division.

- May be there is a story behind extremal auxiliary random variables that will shed light into decades long open problems.

Thank You

CONCLUSION

- Cover introduced the broadcast channel
- In that he talked about the Spanish-English channel (randomized time-division)

- Later Marton and Cover-van der Meulen came up with improved achievable regions using auxiliary random variables

- After trying hard and determining the extremal auxiliary random variables, we see that regions revert to randomized time-division (binary) and an as yet undetermined generalization of randomized time-division.

- May be there is a story behind extremal auxiliary random variables that will shed light into decades long open problems.

Thank You

Conclusion

- Cover introduced the broadcast channel
- In that he talked about the Spanish-English channel (randomized time-division)

- Later Marton and Cover-van der Meulen came up with improved achievable regions using auxiliary random variables

- After trying hard and determining the extremal auxiliary random variables, we see that regions revert to randomized time-division (binary) and an as yet undetermined generalization of randomized time-division.

- May be there is a story behind extremal auxiliary random variables that will shed light into decades long open problems.

Thank You