Generalized Rank Weight of Linear Codes and Its Application to Secure Network Coding

Ryutaroh MATSUMOTO[1] (speaker)
(jont work with Jun KURIHARA[2], and
Tomohiko UYEMATSU[1])

[1] Tokyo Institute of Technology, Japan [2] KDDI Laboratories, Japan

Institute of Mathmatics for Industry Workshop
Kyushu University, Japan
March 2013
(Please ask your question at any time.)

Structure of this talk

1. Review
    - Network coding
    - Secure network coding
    - Silva and Kschischang's secure network coding
2. Definition of new parameters of linear codes and its meaning
3. Relation to known parameters
4. Conclusion
5. Appendix: Proofs of key theorems
6. Appendix: Application to secret sharing
7. Appendix: Extension to non-uniform distributions

Pp.21–32 are not included in the proceedings at your hand.
This slide is available at `http://www.rmatsumoto.org/rgrw.pdf`.

Note that itemizations and enumerations are somehow not printed in the proceedings.

Routing vs. network coding

We consider computer networks and wireless networks with multiple nodes.

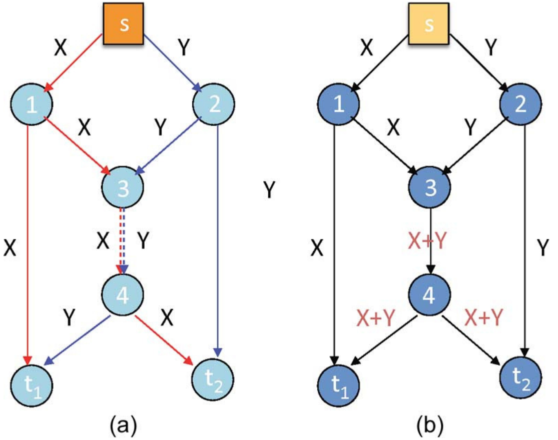
Routing An intermediate node only forwards incoming data.

Network Coding (NC) An intermediate node combines data incoming from its multiple links and jointly encodes them.

Advantages of NC are

- Throughput can increase.
- Optimal NC can be constructed easily while finding optimal routing is NP-hard.

Most famous example of NC — The butterfly network



(a)

(b)

NC multicasts 2 symbols to 2 sinks while routing can multicast at most 1 symbol.

Linear processing works well as the coding in NC.

Consider the single source multicast scenario (All sinks want the same information). The maximum throughput is the minimum of max-flow between the source and each sink.

Secure network coding

Assumptions:

- single source multicast, and
- an adversary (Eve) can eavesdrop her chosen μ links in the network.

Goal: The legitimate users want to hide transmitted data from Eve.
The above problem and its solution were proposed as "secure network coding" by Cai and Yeung (2002).

Relation to other areas:

- Secure network coding is the network coding counterpart of the wiretap channel coding initiated by Wyner (1975) and Csiszár-Körner (1978).
- Secure network coding is a generalization of (threshold-type linear) secret sharing proposed by Shamir and Blakley (1979).

Secure network coding

Questions:

- How does one model Eve's observation?
- How does one measure the amount of information obtained by Eve's observation?

I will review necessary materials to answer the above questions.

Transfer matrix in linear network coding

Assume:

- single source multicast, linear processing at every node,
- network can be modeled as an acyclic graph,
- delay can be ignored, or data generated by the source at the same time are linearly combined,
- the source node has n outgoing links,
- m consecutive time slots are used for the source to send one packet, and
- a link can carry one $GF(q)$ symbol per one time slot,

x_{ij}: $GF(q)$ symbol at time j on the i-th outgoing link from the source.

Fix μ links $e_1, \ldots, e_\mu$. z_{ij}: $GF(q)$ symbol at time j on e_i.

Observation at time j =

$$\begin{pmatrix} z_{1j} \\ \vdots \\ z_{\mu j} \end{pmatrix} = B \begin{pmatrix} x_{1j} \\ \vdots \\ x_{nj} \end{pmatrix}, \quad \text{transfer matrix } B \in GF(q)^{\mu \times n}$$

Information leaked by Eve's observation

Random variable S: (secret) message to legitimate receivers
Random variable Z: Eve's (adversary's) observation

Amount of information leaked in Eve's observation
$= I(S; Z) =$ the mutual information between S and Z.

Special case: $I(S; Z) = 0 \Leftrightarrow S$ and Z are statistically independent
$\Leftrightarrow$ Having Z does not help guessing the value of S.

Silva and Kschischang's universal secure network coding

Silva and Kschischang (2011) proposed network coding that is

1. secure
2. error-correcting,
3. and universal (working well with any transfer matrix)

by using

- $C_2 \subsetneqq C_1 \subseteq GF(q^m)^n$,
- with C_1 and C_2 being MRD (maximum rank distance).

Questions

- What is the security performance and the error correction capability when C_1 or C_2 is not MRD?
- What parameter of C_1 and C_2 exactly expresses the security and the error correction capability?

I will answer those questions.

Review of Silva and Kschischang (2011)

single source multicast (acyclic, delay-free)
n: minimum of max-flows ($\simeq$ # outgoing links from the source)
m: time slots in a packet, m must be $\geq n$ for existence of MRD codes.
One $GF(q)$ symbol is carried on a link per time slot
$GF(q)$-linear coding at all intermediate nodes
$m \times n$ $GF(q)$ symbols in a packet.

$C_2 \subsetneqq C_1 \subseteq GF(q^m)^n$: $GF(q^m)$-linear (MRD) codes
A message is a coset $\vec{a} + C_2 = \{\vec{a} + \vec{x} : \vec{x} \in C_2\} \in C_1/C_2$, for $\vec{a} \in C_1$.
$|\vec{a} + C_2| = |C_2|$ for any $\vec{a}$.

The number of messages is

$$= \frac{|C_1|}{|C_2|} = \frac{q^{m \dim C_1}}{q^{m \dim C_2}} = q^{m(\dim C_1 - \dim C_2)}.$$

Generation of a packet from a given message

$C_2 \subsetneqq C_1 \subseteq GF(q^m)^n$: $GF(q^m)$-linear (MRD) codes
$S \in C_1/C_2$: Given message

1. Randomly choose a vector $\vec{x} = (x_1, \ldots, x_n) \in S \subsetneqq GF(q^m)^n$.
2. Expand $x_i \in GF(q^m)$ into $(x_i^{(1)}, \ldots, x_i^{(m)}) \in GF(q)^m$ by some fixed $GF(q)$-linear basis of $GF(q^m)$,
3. Send $x_i^{(j)}$ on link i at time j.

Generation of $\vec{x}$ from S is called the nested coset coding.

Roles of C_1 and C_2

$C_2 \subsetneq C_1 \subseteq GF(q^m)^n$: $GF(q^m)$-linear (MRD) codes
A message is a coset $\vec{a} + C_2 = \{\vec{a} + \vec{x} : \vec{x} \in C_2\} \in C_1/C_2$.

C_1 realizes the error correction. By not using vectors outside of C_1, error correction becomes feasible. Setting $C_1 = GF(q^m)^n$ turns off the error correction capability.

C_2 realizes the secrecy of the message by randomizing it. Setting $C_2 = \{0\}$ removes the randomization and the secrecy of messages.

The same kind of message randomization is used in the wiretap channel coding and the secret sharing for the same purpose.

q-th power of subspaces (Stichtenoth (1990))

$\vec{x} = (x_1, \ldots, x_n) \in GF(q^m)^n$,
$\vec{x}^q = (x_1^q, \ldots, x_n^q)$.
$V^q = \{\vec{x}^q : \vec{x} \in V\}$ for an $GF(q^m)$-linear subspace V of $GF(q^m)^n$.
V^q is again an $GF(q^m)$-linear subspace despite $\vec{x} \mapsto \vec{x}^q$ is $GF(q^m)$-nonlinear.
$V^* = V + V^q + V^{q^2} + V^{q^3} + \cdots + V^{q^{m-1}}$.
$\Gamma = \{V \subseteq GF(q^m)^n : V \text{ is } GF(q^m)\text{-linear and } V^q = V\}$

1. For an $GF(q^m)$-subspace $V \subseteq GF(q^m)^n$, $V^q = V$ iff V has an $GF(q^m)$-basis written in $GF(q)^n$,
2. V^* is the smallest $GF(q^m)$-space in Γ containing V.

The above were given by Stichtenoth (1990) for studying subfield subcodes.

j-th Relative Generalized Rank Weight (RGRW)

For $C_2 \subsetneq C_1 \subseteq GF(q^m)^n$,

$$\begin{aligned} M_j(C_1, C_2) &= \min\{\dim V : V \in \Gamma, \dim C_1 \cap V - \dim C_2 \cap V \geq j\} \\ &= \min\{\dim V : V \in \Gamma, \dim C_1 \cap V - \dim C_2 \cap V = j\} \end{aligned}$$

Eve creates a network of arbitrary shape and choose arbitrary μ links to observe.

Z: observed information, S: secret message (uniform distribution)

Relation between RGRW and eavesdropped information

$\max I(S;Z)$ in $\log_{q^m} \geq j \Leftrightarrow \mu \geq M_j(C_2^\perp, C_1^\perp)$

The maximum is taken over all shapes of network and all choices of μ links.

Corollary

If $\mu < M_1(C_2^\perp, C_1^\perp)$ then there is no information leakage.

I will explain why “rank” is included in its name.

Review of Gabidulin's rank weight

Recall $C_2 \subsetneqq C_1 \subseteq GF(q^m)^n$
The rank weight is related to error correction for network coding.

The legitimate receiver can correct errors occurred at arbitrary t links $\Leftarrow$ the minimum rank weight of C_1 is $\geq 2t + 1$.

$\vec{x} = (x_1, \ldots, x_n) \in GF(q^m)^n$,
$\langle x_1, \ldots, x_n \rangle = \{\sum_{i=1}^{n} a_i x_i : a_i \in GF(q)\} \subset GF(q^m)$.

$w_R(\vec{x})$ = Gabidulin's rank weight = $\dim_{GF(q)} \langle x_1, \ldots, x_n \rangle$
The minimum rank weight $d_R(C_1)$ of C_1 = $\min\{w_R(\vec{x}) \mid \vec{0} \neq \vec{x} \in C_1\}$.

Relation to Gabidulin's rank weight

Recall $C_2 \subsetneq C_1 \subseteq GF(q^m)^n$
The proposed 1st RGRW $M_1(C_1, C_2)$ is related to w_R as

$$M_1(C_1, C_2) = \min\{w_R(\vec{x}) : \vec{x} \in C_1 \setminus C_2\}.$$

$M_1(C_1, \{\vec{0}\})$ = the minimum rank weight $d_R(C_1)$ of C_1.

RGRW generalizes Gabidulin's rank weight.

Relation to (relative) generalized Hamming weight

$I \subseteq \{1, \ldots, n\}$
$V_I = \{\vec{x} \in GF(q^m)^n : x_i = 0 \text{ if } i \notin I\}$
$\dim V_I = |I|$
$C_1 \cap V_I$ is the shortened code of C_1 to the index set I.

j-th generalized Hamming weight (GHW) of $C_1 \subseteq GF(q^m)^n$
$= \min\{\dim V_I : \dim C_1 \cap V_I \geq j\}$ (V.K. Wei (1991))

j-th relative generalized Hamming weight (RGHW) of $C_2 \subsetneq C_1 \subseteq GF(q^m)^n$
$= \min\{\dim V_I : \dim C_1 \cap V_I - \dim C_2 \cap V_I \geq j\}$ (Luo et al. (2005))

Recall that RGRW was

$$M_j(C_1, C_2) = \min\{\dim V : V \in \Gamma, \dim C_1 \cap V - \dim C_2 \cap V \geq j\}.$$

The difference between RGRW and RGHW is the set of intersecting subspaces. Our naming of RGRW follows the conventions set by Gabidulin and Luo et al.

Error correction and RGRW

Fix single sink with N incoming links.
$A \in GF(q)^{N \times n}$: transfer matrix from the source to the sink
The sink knows A (coherent network error correction)
t broken links inject erroneous symbols from time 1 to m.

The sink can correct any t link errors with any A of rank$A \geq n - \rho$ iff $M_1(C_1, C_2) > 2t + \rho$.

Summary

- j-th relative generalized rank weight $M_j(C_1, C_2)$ was introduced for $C_2 \subsetneqq C_1 \subseteq GF(q^m)^n$.
- $\max I(S; Z) \geq j \Leftrightarrow$ the number of eavesdropped link $\geq M_j(C_2^\perp, C_1^\perp)$.
- $C_2 \subsetneqq C_1$ can correct t link errors and ρ rank deficiency iff $2t + \rho < M_1(C_1, C_2)$.
- RGRW is a generalization of Gabidulin's rank weight and is related to Luo et al.'s relative generalized Hamming weight.

All proofs are available from ~~arXiv:1207.1936 or the final version of the Allerton 2012 conference proceedings~~ arXiv:1301.5482. The GRW (without the 1st R) was also concurrently and independently introduced in F. Oggier and A. Sboui, "On the existence of generalized rank weights," in Proc. ISITA 2012, Honolulu, Hawaii, USA, Oct. 2012, pp. 406–410.

What can be done next?

By this result, we can evaluate the error correction capability and the security of any $C_2 \subsetneq C_1$. I expect that more convenient $C_2 \subsetneq C_1$ exist, e.g. with smaller m of $GF(q^m)$ (for MRD C_1, C_2, m must be $\geq n$).
This direction of research was also pointed out by M. Médard at Q&A of the Allerton conference, but our investigation has not started.

Proof sketch: the relation between secrecy and dimension 1

Random variable $S \in C_1/C_2$: Given message
Random variable $X \in S \subset GF(q^m)^n$: transmitted codeword (or packet)
$B \in GF(q)^{\mu\times n}$: a fixed transfer matrix
Assumption: X and S are uniformly distributed.
$\Rightarrow$ As an RV, X can take any vector in C_1.

$I(BX; S) = H(BX) - H(BX|S)$
The uniformity assumption implies

$$\begin{aligned} H(BX) &= \log_{q^m} \text{ the number of possible } BX \\ &= \log_{q^m} |\text{image of map } X \mapsto BX| \\ &= \dim C_1 - \dim(C_1 \cap \ker(B)) \end{aligned}$$

$\ker(B)$ as a linear map from $GF(q^m)^n$ to $GF(q^m)^\mu$.

Proof sketch: the relation between secrecy and dimension 2

$I(BX;S) = H(BX) - H(BX|S)$
The uniformity assumption also implies

$$\begin{aligned} H(BX|S) &= \log_{q^m} \text{ the number of possible } BX \text{ given } S = s \\ &= \log_{q^m} \text{ the number of possible } BX \text{ given } S = C_2 \\ &= \log_{q^m} |\text{image of map } S \to BS| \\ &= \dim C_2 - \dim(C_2 \cap \ker(B)) \end{aligned}$$

$\Rightarrow I(BX;S) = \dim C_1 - \dim C_2 - (\dim(C_1 \cap \ker(B)) - \dim(C_2 \cap \ker(B)))$.

Extension of Forney's second duality lemma

For any space $V \subset GF(q^m)^n$ we have

$$\dim C_1 \cap V - \dim C_2 \cap V = \dim C_1/C_2 - \dim(C_2^\perp \cap V^\perp) + \dim(C_1^\perp \cap V^\perp).$$

Substituting the above extension of Forney's lemma into

$$I(BX;S) = \dim C_1 - \dim C_2 - (\dim(C_1 \cap \ker(B)) - \dim(C_2 \cap \ker(B)))$$

yields

$$I(BX;S) = \dim(C_2^\perp \cap \ker(B)^\perp) - \dim(C_1^\perp \cap \ker(B)^\perp).$$

Universality leads RGRW

For a fixed $B \in GF(q)^{\mu\times n}$, we have
$I(BX;S) = \dim(C_2^\perp \cap \ker(B)^\perp) - \dim(C_1^\perp \cap \ker(B)^\perp)$.
The universal security deals with all $B \in GF(q)^{\mu\times n}$:

$$\begin{aligned}
& \max_{B\in GF(q)^{\mu\times n}} I(BX;S) \\
= & \max_{B\in GF(q)^{\mu\times n}} \dim(C_2^\perp \cap \ker(B)^\perp) - \dim(C_1^\perp \cap \ker(B)^\perp) \\
= & \max_{V\in\Gamma, \dim V \le \mu} \dim(C_2^\perp \cap V) - \dim(C_1^\perp \cap V) \\
& \text{(by the 1st item in p.13)}
\end{aligned}$$

Recall

$$\begin{aligned}
M_j(C_2^\perp, C_1^\perp) &= \min\{\dim V : V \in \Gamma, \dim C_2^\perp \cap V - \dim C_1^\perp \cap V \ge j\} \\
&= \min\{\dim V : V \in \Gamma, \dim C_2^\perp \cap V - \dim C_1^\perp \cap V = j\}
\end{aligned}$$

Claims in p.14 follow from the above.

Some remarks on the proof argument

- One can remove all the assumptions on the probability distributions of S and X, which make the proof more complicated.
- One can deduce the relation between secret sharing and RGHW by restricting the set of matrices B.

Threshold-type linear secret sharing

Goal: Distribute a secret S to n participants so that

- Any α or more participants can recover S, and
- Any β or less participants have NO information on S.

We need to evaluate α and β.
Coding method:
$C_2 \subset C_1 \subset GF(q)^n$
$S \in C_1/C_2$: Given secret
$S \ni X = (x_1, \ldots, x_n)^T$: randomly chosen vector
x_i is distributed to the i-th participant.

Recoverability

Recovery of S by a subset of n participants is equivalent to the erasure decoding by the subset of participants for C_1/C_2. Recoverability is completely determined by the coset distance (= 1st RGHW) of C_1/C_2 (Duursma and Park 2010).

Information gained by fixed μ participants

$C_2 \subset C_1 \subset GF(q)^n$
$S \in C_1/C_2$: Given secret
$S \ni X = (x_1, \ldots, x_n)$: randomly chosen vector
Fixed μ participants have $(x_{i_1}, \ldots, x_{i_\mu})$.
$B \in GF(q)^{\mu \times n}$ such that $(x_{i_1}, \ldots, x_{i_\mu})^T = BX$. Every entry in B is either 0 or 1.

$I(BX; S) = \dim(C_2^\perp \cap \ker(B)^\perp) - \dim(C_1^\perp \cap \ker(B)^\perp)$.

On the other hand, $\ker(B)^\perp = V_I$ with $I = \{i_1, \ldots, i_\mu\}$.
$\Rightarrow I(BX; S) = \dim(C_2^\perp \cap V_I) - \dim(C_1^\perp \cap V_I)$.

Recall $V_I = \{\vec{x} \in GF(q)^n : x_i = 0 \text{ if } i \notin I\}$.

Worst-case information gain by arbitrary μ participants

$\max_B$ is taken over all possible combinations of μ participants.

$$
\begin{aligned}
& \max_B I(BX;S) \\
= \quad & \max_B \dim(C_2^\perp \cap \ker(B)^\perp) - \dim(C_1^\perp \cap \ker(B)^\perp) \\
= \quad & \max_{|I|\leq\mu} \dim(C_2^\perp \cap V_I) - \dim(C_1^\perp \cap V_I)
\end{aligned}
$$

Recall that j-th RGHW of $C_2^\perp$ and $C_1^\perp$
$= \min\{\dim V_I : \dim C_2^\perp \cap V_I - \dim C_1^\perp \cap V_I \geq j\}$

The above leads to …

Secret sharing and RGHW

Z: shares of arbitrary μ participants, S: secret (uniform distribution)

Relation between RGHW and Z

$\max I(S;Z)$ in $\log_{q^m} \geq j \Leftrightarrow \mu \geq j$-th RGHW of $C_2^\perp$ and $C_1^\perp$

The maximum is taken over all choices of μ participants.

Corollary

If $\mu <$ 1st RGHW (= coset distance) of $C_2^\perp$ and $C_1^\perp$ then there is no information leakage of S into Z.

The above were reported at J. Kurihara et al.,
`http://dx.doi.org/10.1587/transfun.E95.A.2067`

Extension to non-uniform distributions

S: secret message, X: transmitted packet (codeword)
(S, X) are often assumed to be uniformly distributed.
Zhang and Yeung considered artibrary distributions of (S, X) (ISIT 2009).

For extension of our result, evaluation of $I(BX; S)$ for a fixed B is enough. For any $B \in \mathcal{B}$ and $A \in \mathcal{A}(B) \subset \mathcal{A}$,

$$\begin{aligned} H(A) &= \log|\mathcal{A}| - D(A\|U_{\mathcal{A}}) \text{ (see an information theory textbook)}, \\ H(A|B) &= \mathbf{E}_B[\log|\mathcal{A}(B)|] - D(A\|U_{\mathcal{A}(B)}|B) \text{ (similarly shown as above)}. \end{aligned}$$

When $B = b$, possible realizations of A is narrowed to $\mathcal{A}(b) \subset \mathcal{A}$.
$U_{\mathcal{A}(B)}$: RV conditionally uniform on $\mathcal{A}(B)$ given B.
By using the above, ...

Extension to non-uniform distributions (contd.)

$$
\begin{aligned}
H(A) &= \log|\mathcal{A}| - D(A\|U_{\mathcal{A}}),\\
H(A|B) &= \mathbf{E}_B[\log|\mathcal{A}(B)|] - D(A\|U_{\mathcal{A}(B)}|B)
\end{aligned}
$$

give

$$
\begin{aligned}
H(S) &= \dim C_1/C_2 - D(S\|U_{C_1/C_2}),\\
H(X|S) &= \dim C_2 - D(X\|U_S|S),\\
&\vdots
\end{aligned}
$$

By using the above, one has, for a fixed B,

$$
\begin{aligned}
& \dim(C_2^\perp \cap \ker(B)^\perp) - \dim(C_1^\perp \cap \ker(B)^\perp) - D(S\|U_{C_1/C_2})\\
\leq\ & I(BX;S)\\
\leq\ & \dim(C_2^\perp \cap \ker(B)^\perp) - \dim(C_1^\perp \cap \ker(B)^\perp) + D(X\|U_S|S).
\end{aligned}
$$

$D(S\|U_{C_1/C_2})$ quantifies the non-uniformity of S, while
$D(X\|U_S|S)$ quantifies the conditional non-uniformity of X given S.

Extension to non-uniform distributions (contd.)

$$\begin{aligned} & \dim(C_2^\perp \cap \ker(B)^\perp) - \dim(C_1^\perp \cap \ker(B)^\perp) - D(S\|U_{C_1/C_2}) \\ \leq\ & I(BX;S) \\ \leq\ & \dim(C_2^\perp \cap \ker(B)^\perp) - \dim(C_1^\perp \cap \ker(B)^\perp) + D(X\|U_S|S). \end{aligned}$$

$D(S\|U_{C_1/C_2})$ quantifies the non-uniformity of S, while
$D(X\|U_S|S)$ quantifies the conditional non-uniformity of X given S.

Non-uniform S may decrease Eve's information $I(BX;S)$,
while conditionally non-uniform X given S may increase $I(BX;S)$.

One can remove all the assumptions on distributions of S and X in this talk.

All mathematical claims and proofs are available as arXiv:1301.5482.
This slide is available at `http://www.rmatsumoto.org/rgrw.pdf`.