

تئوری رمز‌گذاری و رابطه آن با علم کامپیوتر

حسین صداقتی

چکیده:

با ایجاد کامپیوتر و افزایش تعداد کاربران آن، مسئله رمز‌گذاری و رمزنگاری که از قدیم نیز وجود داشت، جنبه جدید و مهمی را به خود گرفت. رمزنگاری و رمز‌گذاری هنر نوشتن به صورت رمز است، بطوریکه هیچ کس به غیر از دریافت کننده مورد نظر، نتواند محتوای پیغام را بخواند. این کار به خاطر دلایل مختلفی انجام می شود که از مهمترین آن دلایل می توان به امنیت ایجاد شده و حفظ محتوای پیام ها و اطلاعات در هنگام استفاده از آن اشاره کرد. در این مقاله مطالبی را در ارتباط با تئوری کدینگ و انواع روش های آن بیان می کنیم.

کلید واژه ها:

رمز‌گذاری، تئوری رمز‌گذاری، کدگذاری منبع، کدگذاری کانال، تئوری رمز‌گذاری جبری، کد سدکننده خطی، کد حلقه ای، روش آلبرتی، رمز‌گذاری متقارن، رمز‌گذاری نامتقارن، RSA

مقدمه:

رمز‌گذاری یا همان رمزنگاری، یک نوع علم و هنر محسوب می شود. علم است، چون درون آن الگوریتم های زیادی وجود دارد که در بعضی مواقع پیچیده اند و در ثانی هنر است، چون استفاده کردن از الگوریتم ها به نحو مناسب و در جای خود، چیزی کمتر از هنر نیست. رمزنگاری در دنیای تجاری امروز بسیار اهمیت دارد، چون که ساده ترین و کاربردی ترین روش حفاظت از داده هایی است که به صورت الکترونیکی، ذخیره، پردازش و انتقال داده می شوند (آنگوین جولیا، ۲۰۰۰).

مورد استفاده رمز‌گذاری در بسیاری از امور کسب و کار می باشد. برای مثال، این اجازه را به تجار می دهد تا شماره حساب های مشتریان خود را محافظت نمایند و داد و ستد های خود را به خوبی و با اطمینان انجام دهند. حتی در مورد قرارداد های قانونی که باید از طریق اینترنت انتقال داده شوند، رمز‌گذاری، امنیت و حفاظت این قرارداد ها را فراهم می کند. در این مقاله، در آغاز به معرفی تئوری رمز‌گذاری و در ادامه انواع روش های کدگذاری و متد های آن را بررسی می کنیم.

متن مقاله:

رمزگذاری عبارت است از فرایند تغییر شکل اطلاعات الکترونیکی در یک فرم خاص که تنها توسط یک شخص یا عده ای خاص قابل خواندن و ترجمه شدن باشد. تئوری رمزگذاری یا تئوری رمزنگاری (Coding Theory)، یک شاخه از علوم کامپیوتر و ریاضی است که با فرایندهای متمایل به خطا (Error-prone) از طریق انتقال داده ها در کانال های ارتباطی شلوغ کار می کند. بنابراین، تعداد بسیار زیادی از اشتباهات و خطاهایی که در کانال های ارتباطی ایجاد می شوند، قابل تصحیح و درست شدن می باشند. همچنین این تئوری با خصوصیات کدها و رموزها ارتباط برقرار می کند و آن ها را برای کاربرد مناسب خود در جای مناسب راهنمایی و هدایت می کند. دو طبقه بندی از کدها و رموزها وجود دارد:

(۱) کدگذاری منبع (Source Coding یا Entropy Coding)

(۲) کدگذاری کانال (Channel Coding یا Forward Error Correction)

(۱) کدگذاری منبع، تلاش می کند تا داده ها را بصورت فشرده (Compress) از یک منبع در بیاورد تا بتواند آن ها را به صورت اثربخش (Efficient) انتقال دهد. ما این عمل را هر روز در اینترنت، هنگامی که داده ها را فشرده می سازیم و حجم فایل ها را کمتر می کنیم، مشاهده می کنیم. با این کار، بار شبکه یا ترافیک آن (Network Load) را کمتر خواهیم کرد.

(۲) کدگذاری کانال، بیت های داده ای را که بیت های زائد (Redundant bits) نیز خوانده می شوند، برای انتقال داده ها اضافه می کند. با این کار، انتقال داده ها در کانال های ارتباطی با مزاحمت کمتری همراه خواهد شد. برای مثال، یک CD موزیک معمولی را در نظر بگیرید. این CD از یک کد قوی به نام (Reed-Solomon) استفاده می کند که مشکلات و خرابی های روی CD را درست می نماید. در این مثال، کانال ارتباطی همان CD است. مودم های داده ای، وسایل ارتباط تلفنی و حتی NASA (سازمان فضانوردی ایالات متحده آمریکا)، از روش کدگذاری کانال استفاده بسیاری می کنند (www.wikipedia.org).

همانطور که بیان شد، کدگذاری منبع، Entropy Coding نیز خوانده می شود. اما انتروپی چیست؟ انتروپی یک منبع، در واقع اندازه گیری اطلاعات آن منبع می باشد. تکنیک های مختلفی برای کدگذاری منبع وجود دارند که حد انتروپی منبع را تعیین می کنند. برای مثال، اگر $C(X)$ تعداد بیت ریت ها (Bit rates) بعد از فشرده سازی باشد و $H(X)$ انتروپی منبع باشد، آنگاه:

$$C(X) > H(X)$$

بنابراین اطلاعات بیشتری بعد از فشرده سازی انتقال داده می شود.

تئوری کدگذاری جبری (Algebraic Coding Theory)، یک زیر مجموعه از تئوری کدینگ محسوب می شود که خصوصیات کد ها را به صورت عبارات جبری بیان می کند. تئوری کدگذاری جبری، خود به دو دسته کد اصلی تقسیم می شود:

(۱) کد های خطی سدکننده (Linear Block Codes)

(۲) کد های حلقه ای (Convolutional Codes)

تئوری کدگذاری جبری، سه خصوصیت هر کد را بررسی و تجزیه می کند:

(a) طول کلمات رمز (b) تعداد کلی کلمات رمزی معتبر (c) حداقل فاصله بین دو کلمات رمزی معتبر

اکنون مطالبی را در مورد کد های خطی سدکننده بیان خواهیم کرد. این کد ها خاصیت خطی

(Linearity) را دارند. به عبارت دیگر، مجموع هر دو کلمه رمز، یک کلمه رمز جدید می باشد و آن ها مانند سدی هستند که جلوی بیت ها را در منبع می گیرند. البته کد های سدکننده ای هستند که خطی نیستند. اما بسیار سخت است که بخواهیم اثبات کنیم که یک کد بدون این ویژگی (Linearity) یک کد خوب است. هر کد سدکننده خطی با حروف (n,m,d) معرفی می شود که البته (d) مینیمم است. n طول کلمات رمزی است که البته با نشانه هایی (Symbols) همراه است. m تعداد نشانه های منبع است که می تواند یکباره برای کدگذاری استفاده شود و d مینیمم کمترین فاصله برای کدگذاری بین کلمات رمزی است. کد های سدکننده خطی، با مشکل Penny packing گره خورده اند که البته در سالیان اخیر توجه بسیار زیادی به این مشکل شده است. این مشکل در دو بعد توضیح داده می شود. برای مثال در نظر بگیرید که یک مجموعه ای از Pennies را بر روی یک میز قرار داده اید و آن ها را به یکدیگر نزدیک می کنید. نتیجه این کار یک نمونه شش گوشه می شود که همانند یک لانه زنبور است. اما کد های خطی ابعاد بسیار بیشتری را نسبت به این لانه زنبور ایجاد شده، دارا هستند که به راحتی قابل دیدن و بررسی کردن نیستند. برای مثال Golay Code، رمزی است که در ارتباطات فضایی با عمق و برد زیاد با 24 بعد قابل استفاده است و مانند مثال ما دو بعدی نیست.

تئوری رمزگذاری از یک مدل N بعدی استفاده می کند و کارایی این مدل را افزایش می دهد. اما کد های حلقه ای چیستند؟ این کد ها در مودم های صوتی (Voice band modems)، نسخه های 17 ، 32 و 34 استفاده می شوند. از این کد ها در شبکه تلفن های همراه (GSM)، ماهواره ها و دستگاه های ارتباطی نظامی نیز استفاده می شود. به طور معمول، کد های حلقه ای هیچ نوع محافظتی را در برابر اختلالات در مقایسه با کد های خطی سدکننده ارائه نمی

کنند. تنها مزیت آن ها، آسانی استفاده از آن ها در مقایسه با کد های خطی سد کننده است. در کد های حلقه ای، رمز کننده (Encoder) معمولاً دارای یک چرخه بسیار ساده است که یک

حافظه (Memory) و تعدادی حلقه بازخورد را داراست. رمز گشا (Decoder) هم می تواند به حالت نرم افزاری یا سخت افزاری نصب و اجرا شود (www.wikipedia.org).

اکنون تعدادی از کاربرد های تئوری رمزگذاری را بیان خواهیم کرد. یکی از کاربرد های این تئوری این است که در طراحی کد هایی که موجب هم زمانی (Synchronization) می شوند، نقش دارد. یک کد می تواند طوری طراحی شود که در فاز تغییر خود به آسانی شناسایی و در صورت داشتن مشکل تصحیح شود. همچنین این کد می تواند سیگنال های چندگانه را از روی یک کانال به طور همزمان ارسال کند. یکی دیگر از کاربرد های تئوری رمزگذاری در تلفن های همراه است. کد هایی به نام Code Division Multiple Access (CDMA)، در این تلفن ها وجود دارند. در واقع هر تلفن یک کلمه رمزی را از یک طبقه خاص (فیلد جبری) اختیار می کند. وقتی که عمل انتقال امواج صورت می گیرد، این کد ها امکان اختلال را در تماس کاهش می دهند و در صورت بروز مشکل، آن را اصلاح می کنند. دسته معروف دیگری از کد ها وجود دارد که به Automatic Repeat reQuest (ARQ) معروف می باشد. این کد ها، بیت هایی را به پیام فرستاده شده اضافه می کنند و آن را طولانی تر می سازند. در واقع، یک سرآیند (Header) را که دارای یک شماره سریال شناسایی است به پیام اضافه می کنند. با این کار، دستگاه گیرنده (Receiver) آن پیام را شناسایی کرده و در صورت عدم تطابق از فرستنده می خواهد تا پیام را دوباره ارسال کند. اکثر شبکه های گسترده (Wide Area Network یا WAN) و پروتکل ها از روش رمزی ARQ استفاده می کنند. پروتکل های معروف SDLC (IBM)، TCP (Internet) و X.۲۵ (International) از جمله پروتکل هایی هستند که از این روش استفاده می کنند (www.wikipedia.org).

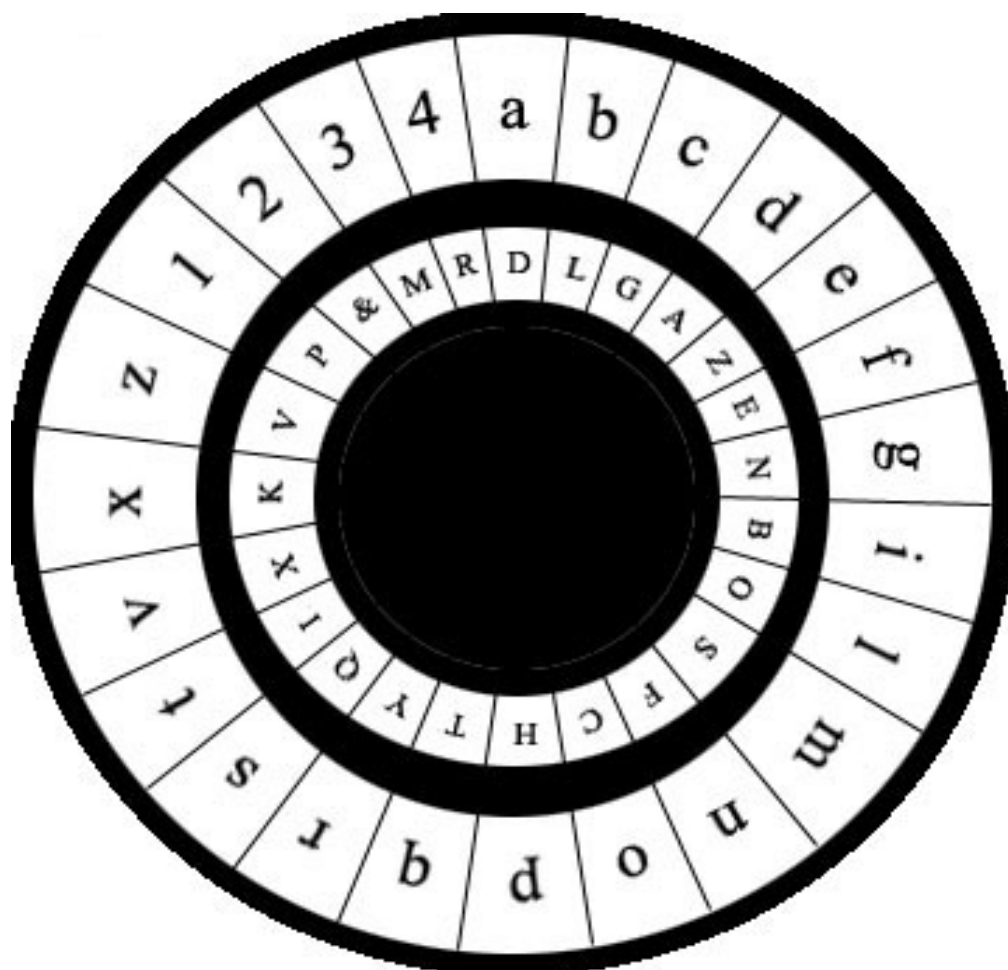
در ادامه بحث، مطالبی را به اختصار در مورد انواع روش های ابداعی کدگذاری بیان می کنیم.

روش آلبرتی:

لئون باتیستا آلبرتی یکی از انواع روش های رمزنگاری را در سال ۱۴۶۶ اختراع کرد که روش او را پایه و اساس علم رمزشناسی نیز می دانند. در واقع او را پدر علم رمزشناسی غرب می دانند. در روش آلبرتی به این ترتیب عمل می شود که او از یک شکل حلقه مانند (Disc) برای کدگذاری استفاده می کند. در قسمت داخلی این دیسک، نشانه هایی وجود دارد که با یک خط به قسمت بیرونی آن دیسک که نشان دهنده یک حرف یا عدد است، مرتبط می شود. بنابراین

شما تنها نیاز دارید که حرف مناسب را بدانید و رمز را تشخیص دهید. در واقع یک عمل تبدیل (Conversion) انجام می شود.

در زیر، شکل دیسک آلبرتی را مشاهده می کنید: (www.trincoll.edu)



در قرن شانزدهم، جیوانی باتیستا پورتا از یک سیستم کلمات کلیدی که نسخه ای از روش آلبرتی بود، استفاده کرد. در این روش یک کلمه کلیدی با استفاده از حروف الفبا مشخص می شود و کلمه کلیدی مجموعه ای از حروف الفبا را در یک سری شامل می شود.

نمونه ای از این روش را در زیر می بینید:

Equivalent to the outer disc

a b c d e f g h i j k l m n o p q r s t u v w x y z
· A L B E R T I C P H D F G H J K M N O S U V W X Y Z
۱ Z A L B E R T I C P H D F G H J K M N O S U V W X Y
N ۲ Y Z A L B E R T I C P H D F G H J K M N O S U V W X
u ۳ X Y Z A L B E R T I C P H D F G H J K M N O S U V W
m ۴ W X Y Z A L B E R T I C P H D F G H J K M N O S U V
b ۵ V W X Y Z A L B E R T I C P H D F G H J K M N O S U
e ۶ U V W X Y Z A L B E R T I C P H D F G H J K M N O S
r ۷ S U V W X Y Z A L B E R T I C P H D F G H J K M N O
۸ O S U V W X Y Z A L B E R T I C P H D F G H J K M N
o ۹ N O S U V W X Y Z A L B E R T I C P H D F G H J K M
f ۱۰ M N O S U V W X Y Z A L B E R T I C P H D F G H J K
۱۱ K M N O S U V W X Y Z A L B E R T I C P H D F G H J
S ۱۲ J K M N O S U V W X Y Z A L B E R T I C P H D F G H
h ۱۳ H J K M N O S U V W X Y Z A L B E R T I C P H D F G
i ۱۴ G H J K M N O S U V W X Y Z A L B E R T I C P H D F
f ۱۵ F G H J K M N O S U V W X Y Z A L B E R T I C P H D
t ۱۶ D F G H J K M N O S U V W X Y Z A L B E R T I C P H
s ۱۷ H D F G H J K M N O S U V W X Y Z A L B E R T I C P
۱۸ P H D F G H J K M N O S U V W X Y Z A L B E R T I C
۱۹ C P H D F G H J K M N O S U V W X Y Z A L B E R T I
۲۰ I C P H D F G H J K M N O S U V W X Y Z A L B E R T
۲۱ T I C P H D F G H J K M N O S U V W X Y Z A L B E R
۲۲ R T I C P H D F G H J K M N O S U V W X Y Z A L B E
۲۳ E R T I C P H D F G H J K M N O S U V W X Y Z A L B
۲۴ B E R T I C P H D F G H J K M N O S U V W X Y Z A L
۲۵ L B E R T I C P H D F G H J K M N O S U V W X Y Z A

در این حالت، حروف بزرگ معادل الفبای قسمت درونی دیسک آلبرتی هستند و حروفی که در بالا هستند، با قسمت بیرونی دیسک آلبرتی معادل هستند. بنابراین، حروف موجود در اولین ردیف جدول با حروفی که در اولین سطر جدول هستند، تشکیل رمز را می دهند که البته باید معادل آن حروف را دو طرف (فرستنده و گیرنده) بدانند و بدین ترتیب آن را رمزگشایی نمایند (ابراهیم الکدی، ۱۹۹۲، ص ۹۷-۹۹).

از انواع روش های دیگر رمزگذاری می توان به روش های رمزگذاری متقارن و نامتقارن اشاره کرد.

در رمزگذاری متقارن (Symmetric Coding)، طرفین ارتباط از یک کلید رمز استفاده می کنند که بایستی نزد آنان مخفی بماند. پرستفاده ترین الگوریتم در این روش، استاندارد رمزنگاری داده Data Encryption Standard (DES) نامیده می شود. DES یک الگوریتم است که از یک کلید ۷ کاراکتری و از دو اصل جایگزینی و تغییر مکان دیتا استفاده می کند. روش دیگر، رمزگذاری نامتقارن (Asymmetric Coding) است. در این روش از دو کلید عمومی و اختصاصی استفاده می شود. کلید خصوصی نزد هر کدام از کاربران و برای رمزگشایی می ماند و کلید عمومی در اختیار عموم قرار می گیرد، که صرفاً برای رمزکردن است (مسعود حجاریان، ۲۰۰۵).

اما RSA چیست؟ این الگوریتم در سال ۱۹۷۷، توسط Ron Rivest، Adi shamir و Leonard Adleman در موسسه تکنولوژی ماساچوست (MIT) معرفی شد (رایوست، شامیر، آدلمن، ۱۹۷۸). RSA یکی از الگوریتم های رمزنگاری نامتقارن است که بر مبنای روابط ریاضی میان اعداد اول، پایه ریزی شده است. رمز عمومی آن، حاصل ضرب دو عدد اول بزرگ است و رمز خصوصی آن از یک مدل ریاضی بین این دو عدد بدست می آید. برای اطمینان نیز می توان از ترکیب RSA و DES (که در مورد آن در قسمت رمزگذاری متقارن توضیح داده شده است) استفاده کرد. حال به بررسی رابطه تئوری رمزگذاری با علم کامپیوتر می پردازیم.

دانشمندان علوم کامپیوتر از اواخر دهه ۱۹۸۰ شروع به تمرکز کردن بر روی کد ها و طراحی الگوریتم های کد کننده کردند. در واقع آن ها با تمرکز خود بر روی تئوری رمزگذاری موجب ایجاد الگوریتم های رمزگشا و رمزنگار کاملتر، درستتر و سریعتر شدند. در سال های اخیر، تکنیک های جدیدی مانند کد های تصحیح کننده خطا (Error Correcting Codes) ایجاد شده اند که به عنوان ابزار های قوی برای حل مسائل علم کامپیوتر استفاده می شوند.

اما الگوریتم های کارایی که برای تئوری رمزگذاری استفاده می شوند، چیستند؟ نتایجی که در سال های اخیر بدست آمده است حاکی از این مطلب است، که یک نقش جدید برای ایجاد ارتباطات قابل اعتماد و امن در کانال های دارای اختلال ایجاد شده است. این مطالب توسط Claud Shannon در سال ۱۹۴۰ معرفی شدند (سارا رایینسون، ۲۰۰۳). یک ریاضی دان دیگر به نام ریچارد هامینگ، یک روش ترکیبی را از کد های تصحیح کننده خطا، که جزئی از تئوری رمزگذاری هستند، در همان زمان ایجاد کرد. یک کد تصحیح کننده خطا، راهی است که بوسیله آن می توان نیم با اضافه کردن داده های زائد به اطلاعات اصلی، خطا ها را در هنگام انتقال تصحیح کنیم. اگر مقدار اطلاعاتی را که کد ها در

واحد زمان و بر حسب بیت از کانال عبور می دهند را با R نشان دهیم، آنگاه می توان گفت که هر کانال دارای یک ظرفیت C است که بیشترین مقدار نرخ اطلاعاتی را که می تواند انتقال دهد، مشخص می کند. کدهای تصحیح کننده خطا به عنوان مجموعه ای از نوارهایی با طول n هستند که به آن ها کلمه کلیدی نیز گفته می شود. بنابراین، نرخ R با گرفتن لگاریتم قدر مطلق ظرفیت (C) تقسیم بر n حاصل می شود. در سال ۱۹۶۰، اولین مقاله توسط آقای رابرت گالاگر در موسسه تکنولوژی ماساچوست ایجاد شد که در آن، رابطه علم کامپیوتر با تئوری رمزگذاری مورد بررسی قرار گرفته است. اما تنها با انقلاب تکنولوژیکی در دهه ۱۹۹۰ بود، که مبحث رابطه علم کامپیوتر با تئوری رمزگذاری قوت گرفت. از آن زمان تاکنون، الگوریتم های بسیار زیادی در رابطه با علم کامپیوتر و تئوری رمزگذاری ایجاد شده است که اغلب آن ها کارایی زیادی را هم دارند (سارا رایبسون، ۲۰۰۳).

نتیجه گیری:

در این مقاله به این مطلب پی بردیم که رمزگذاری و رمزگشایی، مفاهیم و عملیاتی بسیار مهم در ارتباطات دنیای امروز ما هستند. مهمترین دلیل استفاده از رمزگذاری این است که می توانیم بوسیله آن امنیت را ایجاد کنیم. این امنیت برای جلوگیری از سرقت و تحریف اطلاعات قابل انتقال در کانال های ارتباطی است. دلایل دیگری هم که می توانیم ذکر کنیم از این قبیلند:

(۱) جلوگیری از ایجاد خطا، (۲) تصحیح کردن خطای ایجاد شده و (۳) ایجاد همزمانی در شناخت خطا و تصحیح آن (Synchronization). همچنین از رمزگذاری در شبکه GSM تلفن همراه نیز استفاده می شود. در ضمن پی بردیم که علم کامپیوتر و تئوری رمزنگاری رابطه بسیار نزدیکی را با یکدیگر دارند و این رابطه با انقلاب تکنولوژیکی در دهه ۱۹۹۰، خود را بیشتر نشان داد. انواع روش ها و الگوریتم های ارتباطی و رمزنگاری را نیز مانند روش آلبرتی، رمزگذاری متقارن، رمزگذاری نامتقارن و RSA را نیز بیان کردیم. می توان گفت که با پیچیده شدن روزافزون علوم (نه تنها علوم کامپیوتری)، نقش رمزگذاری و رمزنگاری پیش از پیش اهمیت پیدا خواهد کرد.

منابع:

منبع فارسی:

حجاریان، مسعود، کلیات، فایل پی دی اف فناوری اطلاعات، ۲۰۰۵.

منابع انگلیسی:

- ۱) Angwin, Julia. "Internet Encryption 's password is 'slow' ". Wall street Journal. March ۲۸, ۲۰۰۰.
- ۲) Ibrahim A.Al-kadi. "The origins of Cryptology: Arab Contributions", Cryptologia, ۱۶(۲), ۱۹۹۲.
- ۳) R.Rivest, A.Shamir, L.Adleman. "A method for obtaining digital Signatures and public key crypto systems". ACM, ۱۹۷۸.
- ۴) Robinson, Sara. "Coding Theory meets Theoretical Computer Science". SIAM News, Vol.۳۴,۲۰۰۳.

منابع اینترنتی:

- ۱) <http://www.wikipedia.org/wiki/coding.html> [Accessed ۳۱/۱/۲۰۰۷].
- ۲) <http://www.trincoll.edu> [Accessed ۱/۲/۲۰۰۷].