第 55 卷 第 8 期 2006 年 8 月
1000-3290/2006/55(08)/3964-05

物 理 学 报
ACTA PHYSICA SINICA

Vol.55,No.8,August,2006

一种新型的混沌伪随机数发生器*

王 蕾 汪芙平 王赞基

(清华大学电机工程与应用电子技术系,电力系统国家重点实验室,北京 100084)

(2005 年 11 月 23 日收到;2006 年 4 月 10 日收到修改稿)

针对 z-logistic 这类特殊的混沌映射,实现了有限位计算精度下其真实演化轨道的精确计算.将该生成轨道的二值粗粒化输出用作伪随机序列,很大程度上保留了定义在实数域上混沌随机数发生器作为理想信息源的统计特性和随机特性,使得这种伪随机数发生器优良的统计分布和密码学性能得到理论上的强力支持.此外,该伪随机数发生器的周期长度可准确预测,采用简单算法可有效排除产生短周期的弱密钥,克服了传统混沌伪随机数发生器存在弱密钥且无法简单排除的重大缺陷.理论分析和数值实验验证了这种新型混沌伪随机数发生器在周期长度、统计分布和密钥安全性上良好的性能,表明它在众多应用领域(包括数据加密)中具有潜在的应用前景.

关键词:混沌,伪随机数发生器,信息源

PACC:0545

1. 引 言

伪随机数发生器在许多科学技术和工程领域(如密码学、蒙特卡罗仿真和扩频通信)中有着十分广泛的应用.混沌运动是一类极其特殊的运动形式,它遵循确定性动力机制,但表现内在的随机性,因而非常适合用于产生伪随机数.近些年来,应用混沌的良好特性构造伪随机数发生器,即基于混沌的伪随机数发生器(CPRNG),已成为一个研究的热点[1—13].研究表明[2—4],当混沌映射及其输出函数满足特定的约束时,混沌轨道的粗粒化输出序列将成为严格的马尔可夫随机信息源,基于这一粗粒化过程构造的伪随机数发生器在理论上具有很高的安全性.当前这类 CPRNG 主要由数字运算器件完成,定义在实数域上的混沌映射以浮点(或定点)方式通过数值计算实现,本文将这种基于有限精度数值计算的算法称为传统的 CPRNG.传统的 CPRNG 在应用中遇到一个严重的问题,由于有限精度舍入误差的不断累积,在某些种子(混沌轨道初值)下输出序列只有很短的周期,并且没有简单算法事先排除这些种子[14,15].这种 CPRNG 用于构造序列密码算法[6]和分组密码算法[16]时,由于弱密钥(混沌轨道的初值和混沌映射的某些参数用作加密算法的密钥)的存在,它的安全性受到很大质疑.

为克服传统 CPRNG 这一重大缺陷,本文提出了一种新型的 CPRNG 实现方案,针对一类特殊的称为 z-logistic 的混沌映射[17],根据这类映射在有理数支撑域上的等价定义形式,实现了在有限运算精度下计算该支撑域上指定初值的精确轨道,避免了舍入误差的累积并能够借助数论的知识有效排除弱的密钥,从而保留了 CPRNG 理论上的安全性.本文将简述混沌随机数发生器的原理并简单分析传统 CPRNG 安全性的困难,给出新型 CPRNG 的具体形式,详细讨论该 CPRNG 良好的统计性能和安全性能.

2. 混沌运动的信息源特性

以离散时间混沌动力系统为例,利用混沌运动构造信息源的典型形式为

$$x_{n+1} = f(x_n), \tag{1}$$

$$X_n = C(x_n), \tag{2}$$

式中,$\{x_n\}$为混沌轨道,$f: D \to D$,$D \subset \mathbb{R}^m$为混沌动力机制,$C: D \to \{0,1,\cdots,M\}$为粗粒化输出函数,$\{X_n\}$为输出的 M 元序列.Kohda 等[2,4]针对一维混沌系统,通过对混沌映射和输出函数施加一些约束,得到由混沌映射构造同分布离散无记忆信息源的一个充分条件.

* 教育部科学技术研究重大项目(批准号:105004)资助的课题.

命题 1[2]（混沌构造随机数发生器的充分条件）对于(1),(2)式的混沌动力系统和输出函数，如果它们满足下面三个条件，则生成的符号序列 $\{X_n\}$ 为同分布离散无记忆信息源.

(ⅰ)混沌映射 $f:D=[d,e]\to D$ 为一维分段单调满射，即存在 D 的一个划分，$d=d_0<d_1<\cdots<d_M=e$，使得对任意 $i=1,\cdots,M(M\geqslant 2)$，$f$ 在 $D_i=[d_{i-1},d_i]$ 上的限定映射 $f_i(x)$ 为二次可微函数，满足 $f(D_i)=D(i=1,\cdots,M)$，且 f 具有唯一的连续不变测度，记为 $\mu(x)\mathrm{d}x$.

(ⅱ)f 满足等概率分布特性，即

$$|g_i'(x)|\mu(g_i(x))=\frac{\mu(x)}{M}\quad(1\leqslant i\leqslant M),\tag{3}$$

式中 $g_i(x)=f_i^{-1}(x)$ 为 x 的第 i 个原象.

(ⅲ)输出函数 $C(\cdot)$ 满足恒和属性，即

$$\frac{1}{M}\sum_{i=1}^{M}C(g_i(x))=\int_{x\in D}C(x)\mu(x)\mathrm{d}x.\tag{4}$$

对命题 1 所给的充分条件，如果再添加一个约束，即

$$\int_{x\in D_i}\mu(x)\mathrm{d}x=\frac{1}{M}\quad(i=1,\cdots,M),\tag{5}$$

则 $\{X_n\}$ 为一个 M 进制的理想信息源(等概率离散无记忆信息源).

传统的混沌伪随机数发生器以(1)式中混沌轨道的初值 x_0 作为种子(密钥)，通过数值计算在有限精度下实现混沌映射的迭代.由于混沌映射 f 定义在实数域内，有限精度的数值计算必然引入舍入误差，该数值计算过程可表示为

$$\tilde{x}_{n+1}=Q\circ f(\tilde{x}_n)=f(\tilde{x}_n)+\varepsilon_n,\tag{6}$$

式中，$\{\tilde{x}_n\}$ 为有限精度下迭代生成轨道，$Q(\cdot)$ 为量化函数，ε_n 为舍入误差.混沌对初值的敏感依赖性决定了 $\{\tilde{x}_n\}$ 并不是一条真实的混沌轨道.无论 ε_n 有多小，随着演化的进行 $\{\tilde{x}_n\}$ 与真实轨道 $\{x_n\}$ 将完全不相关，它们输出序列的统计特性表现较大差异，基于混沌的相关理论并不适合对这类 CPRNG 的统计特性和安全性能进行分析.

相关研究表明[14,15]，这种基于有限精度数值计算的 CPRNG 在密码学意义上是不安全的，主要表现在以下三个方面：(ⅰ)(6)式的迭代过程本质上是一个有限状态机，必然生成周期轨道，但由于演化机制 $Q\circ f(\cdot)$ 相当复杂，周期轨道的长度不能简单预测.(ⅱ)存在某些初值 x_0，从它们出发的离散化轨道将退化为很短的周期轨道.这些初值应视为弱的密钥，它们极大威胁密码算法的安全.尽管人们提出了许多克服短周期的办法，如采用 m 伪随机序列加扰[18]、使用高维的混沌系统等，这些方法从本质上是通过提高动力系统的维数以增强动力系统的复杂性，从而增大输出序列的总体平均周期，但在具体实现中都不可避免遭受有限精度的困扰，无法有效消除弱密钥.(ⅲ)由于周期轨道长度预测上的困难，目前除了穷举之外，没有更有效的方法找到这些弱密钥并加以排除.

3. 一种新型的 CPRNG

文献[19]提出了一种精确计算 logistic 混沌轨道的实用算法，能够克服迭代过程中有限精度造成的误差累积.本文将该算法进行扩展，针对 z-logistic 混沌映射，提出一种新型的 CPRNG 实现方案.我们选用的 z-logistic 混沌动力机制及输出函数为

$$x_{n+1}=f(x_n)=\sin^2(z\arcsin\sqrt{x_n})\quad(x_n\in(0,1)),\tag{7}$$

$$X_n=C(x_n)=\begin{cases}0 & (x_n<0.5),\\ 1 & (x_n>0.5),\end{cases}\tag{8}$$

式中 z 为混沌映射的参数，为正的偶数，系统的 Lyapunov 指数为 $\lambda=\log_2 z$.容易验证(7),(8)式满足命题 1 的充分条件和附加约束，因此 $\{X_n\}$ 在理论上可以成为理想的信息源.基于该混沌动力机制的 CPRNG 的具体算法叙述如下：

(ⅰ)密钥(种子)为 (m,z,l_0)，m 为大的素数，z 为 $\mathbb{Z}_m^*$（与 m 互素的所有整数的集合）的一个生成元，l_0 为满足 $1\leqslant l_0\leqslant m-1$ 的任意整数.

(ⅱ)混沌迭代过程.令 $n=1,2,\cdots$，重复以下操作：

$$l_n=zl_{n-1}(\mathrm{mod}\,m),\tag{9}$$

$$x_n=\sin^2(l_n\pi/m).\tag{10}$$

(ⅲ)按(8)式输出二进制伪随机序列 $\{X_n\}$.

以上混沌迭代过程利用了这样一个事实：(7)式的混沌映射在同胚映射 $x=\sin^2(\pi c)$ 作用下与映射 $c_{n+1}=zc_n(\mathrm{mod}1)$，$c_n\in(0,1)$ 拓扑共轭，因此 z-logistic 混沌轨道可由(9),(10)式生成.可对于任何形如 $x_0=\sin^2(c_0\pi)$，$c_0\in Q$（有理数域）的初值，通过令 $c_0=l_0/m$，l_0 和 m 为互素的正整数，(9)式将生成精确的轨道 $\{l_n\}$.设函数 $\sin^2(\cdot)$ 由计算机实现时所引入的最大截断误差为 δ，则在有限精度下(9),(10)式

实际生成的混沌轨道 $\{\tilde{x}_n\}$ 与真实轨道 $\{x_n\}$ 之间满足

$$|x_n - \tilde{x}_n| < \delta. \tag{11}$$

在计算精度足够高时，$\delta \ll 1$. 因此，$\{\tilde{x}_n\}$ 与真实轨道 $\{x_n\}$ 将输出几乎完全相同的伪随机序列.

4. 性能分析

由于这种新型 CPRNG 是基于混沌动力机制作用下一条真实运动轨道的输出，因此它的性能可以依据混沌的相关理论定性分析. 下面就伪随机数发生器最重要的周期性能、统计性能和用于数据加密时的密钥安全性能进行讨论，并配合仿真实验加以验证.

4.1. 周期特性

(9)，(10)式的迭代过程等价于将生成轨道限定在有理数域上，因此该轨道必然是周期的，它是混沌动力系统的一条不稳定周期轨道. 在选取密钥 (m, z, l_0) 时，要求 z 为 $\mathbb{Z}_m^*$ 的一个生成元，$\mathbb{Z}_m^*$ 为与 m 互素的所有整数的集合，具体算法见文献[20]. 由数论的相关知识[20]知，在该密钥选取规则下，从 $x_0 = \sin^2(l_0\pi/m)$ 出发的混沌不稳定周期轨道长度为 $N = (m-1)/2$，且与 l_0 取值无关. 因此在取 m 为大素数的情况下，该 CPRNG 不会出现短周期现象，有效排除了弱的密钥(种子).

4.2. 统计特性

由混沌相关理论[21]知，混沌吸引子内的不稳定周期轨道是稠密的，混沌吸引子的遍历特性保证了在不稳定周期轨道足够长的情况下，该周期轨道的统计特性逼近混沌吸引子的统计特性. 因此在素数 m 足够大的情况下，这种新型的 CPRNG 生成的混沌不稳定周期轨道足够长，其输出伪随机序列极大程度保留了命题 1 叙述的混沌随机数发生器的信息源特性.

为验证生成伪随机序列良好的统计特性，这里采用卡方检测[22]的方法对三组密钥下生成序列的随机性进行检测. 这是一种假设检验的方法，如果序列是随机的，那么统计量 χ^2 应服从卡方分布. 该统计量定义为

$$\chi^2 = \frac{2^i}{N}\sum_{s=0}^{2^i-1} Y_s^2 - N, \tag{12}$$

式中，N 表示将待检测序列依次以每 i 比特作为一个模块分段得到的模块数，Y_s 表示数值为 s 的模块总数，卡方分布的自由度为 $2^i - 1$. 对于每一组生成序列由(12)式计算 χ^2 的样本观测值 χ^2_{obv}. 设显著性水平 $\alpha = 0.01$，确定相应的临界值 δ_α，使之满足

$$P(\chi^2 > \delta_\alpha) = \alpha, \tag{13}$$

式中 $P(\cdot)$ 表示随机事件发生的概率. 由于 α 的值很小，根据小概率事件的实际不可能性原理，如果 $\chi^2_{\text{obv}} > \delta_\alpha$，则认为被检测的序列不满足随机性假设；如果 $\chi^2_{\text{obv}} < \delta_\alpha$，则表明该序列具有一定的随机性. 表 1 给出密钥分别为

$$K_1 = \{z_1 = 2, l_0/m = 23156/99999787\},$$
$$K_2 = \{z_1 = 2, l_0/m = 31420/99999821\},$$
$$K_3 = \{z_1 = 2, l_0/m = 135724/99999643\}$$

时统计检测的结果，统计检测表明由该方案生成的伪随机序列具有良好统计特性.

表 1 混沌伪随机序列的卡方检测结果

i	自由度	δ_α	$\chi^2_{\text{obv}}(K_1)$	$\chi^2_{\text{obv}}(K_2)$	$\chi^2_{\text{obv}}(K_3)$
1	1	6.6349	0.0484	1.6384	1.2100
2	3	11.345	1.2952	0.3064	0.6520
3	7	18.475	8.7408	3.9936	7.4272
4	15	30.578	11.408	20.304	8.5632
5	31	52.191	26.272	37.254	19.968
6	63	92.010	67.917	83.867	58.534
7	127	166.99	102.86	122.85	117.48

4.3. 密钥的安全性

对于一个安全的密码算法，它的密钥应满足以下两个条件.

(ⅰ)密钥空间应该足够大，对于本文提出的 CPRNG，密钥 m 通常选取为一个大的素数，l_0 可以取 $[1, m-1]$ 上的任意整数，z 可以取 Z_m^* 的任一生成元. 如果实现精度为 L 位，由素数定理可知，$[2^L, 2^{L+1}]$ 之间的素数的个数约为

$$\Phi(m) \approx \frac{2^{L+1}}{\ln 2^{L+1}} - \frac{2^L}{\ln 2^L} = \frac{(L-1)2^L}{L(L+1)\ln 2}. \tag{14}$$

由此可见，密钥空间大小随着密钥长度呈指数增长.

(ⅱ)密文对密钥敏感依赖，通过密文的相关性无法获取密钥的任何信息. 记 (m, z, l_0) 为真实的密

钥，(m', z', l'_0)为攻击者猜测的密钥，$\{x'_n\}$和$\{x_n\}$为对应的密文.对于本文提出的新方案，可分以下四种情况讨论密文和密钥的关系，图 1 分别给出了四种情况下典型的密文相关曲线.

(1)当 $z' \neq z, l'_0/m' = l_0/m$ 时，$\{x'_n\}$和$\{x_n\}$为初值相同的两个混沌系统生成的序列，它们是不相关的.

(2)当 $z' = z, l'_0/m' \neq l_0/m$ 时，$\{x'_n\}$和$\{x_n\}$为相同的混沌系统从不同初值出发得到的混沌序列，由于混沌对初值的敏感依赖，两序列是不相关的.

(3)当 $z' \neq z, l'_0/m' \neq l_0/m$ 时，$\{x'_n\}$和$\{x_n\}$为不相同的混沌系统从不同初值出发得到的混沌序列，两个序列也是不相关的.

(4)当 $z' = z, l'_0/m' = l_0/m$ 时，密码算法被破解.

因此，从输出序列统计特性上不能获得密钥的有用信息.由图 1 可以看出，除密钥被精确猜出外，其余三种情况下密文相关曲线完全相同.

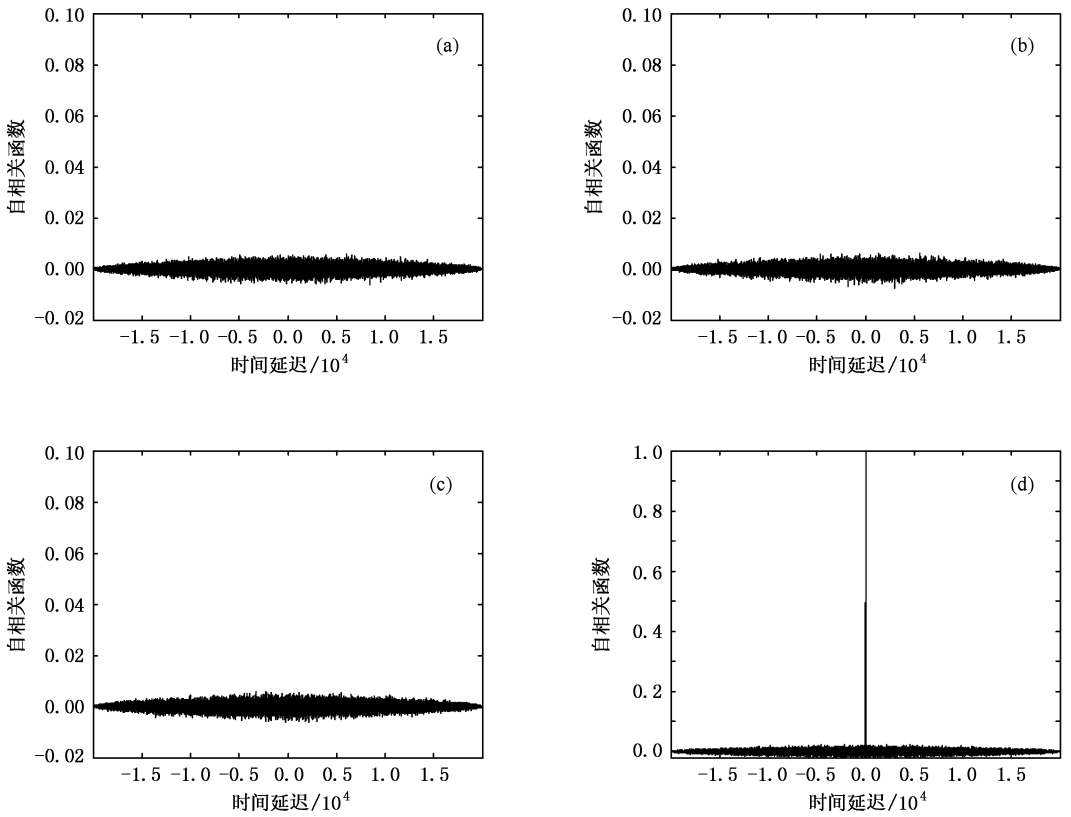

图 1　不同密钥生成密文的相关性　(a)$(z', m'_0, l'_0) = (3, 99999787, 2315687), (z, m_0, l_0) = (2, 99999787, 2315687)$；(b)$(z', m'_0, l'_0) = (2, 99999787, 2315687), (z, m_0, l_0) = (2, 99999821, 3142000)$；(c)$(z', m'_0, l'_0) = (3, 99999787, 2315687), (z, m_0, l_0) = (2, 99999821, 3142000)$；(d)$(z', m'_0, l'_0) = (2, 99999787, 2315687), (z', m'_0, l'_0) = (2, 99999787, 2315687)$

5. 结　论

本文提出的这种新型 CPRNG 本质上是混沌动力机制作用下一条精确不稳定周期轨道的粗粒化输出.由于选用的 z-logistic 映射及其粗粒化输出函数满足生成理想信息源的充分条件，当不稳定周期轨道足够长时，混沌吸引子的统计理论能够保证所生成的伪随机序列很大程度保留了混沌生成理想信息源的优良统计特性和保密性能.由于该 CPRNG 的周期长度可以通过简单算法恰当选取密钥事先确定，克服了传统伪随机数发生器存在弱密钥，且无法简单排除的重大缺陷.仿真实验验证了该新型 CPRNG 优良的周期特性、统计特性和密钥安全性，表明它不仅可以直接应用于蒙特卡罗仿真和扩频通信等领域，而且在安全性能要求很高的数据加密领域也有潜在的应用前景.

[1] Kocarev L 2001 *IEEE Circ. Sys. Mag.* **64** 6
[2] Kohda T 2002 *Proc. IEEE* **90** 641
[3] Sang T, Wang R, Yan Y 2001 *IEEE Trans. Commun.* **49** 620
[4] Kohda T, Tsuneda A 1997 *IEEE Trans. Inform. Theory* **43** 104
[5] Stojanovski T, Kocarev L 2001 *IEEE Trans. Circ. Syst.* Ⅰ **48** 281
[6] Zhou H, Ling X T 1997 *Int. J. Bifurc. Chaos* **7** 205
[7] Xiao F H, Yan G R, Han Y H 2004 *Acta Phys. Sin.* **53** 2877 (in Chinese) [肖方红、阎桂荣、韩宇航 2004 物理学报 **53** 2877]
[8] Hu H P, Liu S H, Wang Z X *et al* 2004 *Acta Math. Phys.* **24** 251 (in Chinese) [胡汉平、刘双红、王祖喜等 2004 数学物理学报 **24** 251]
[9] Cai J P, Li Z, Song W T 2003 *Acta Phys. Sin.* **52** 1871 (in Chinese) [蔡觉平、李 赞、宋文涛 2003 物理学报 **52** 1871]
[10] Sheng L Y, Cao L L, Sun K H *et al* 2005 *Acta Phys. Sin.* **54** 4031 (in Chinese) [盛利元、曹莉凌、孙克辉等 2005 物理学报 **54** 4031]
[11] Peng F, Qiu S S, Long M 2005 *Acta Phys. Sin.* **54** 4562 (in Chinese) [彭 飞、丘水生、龙 敏 2005 物理学报 **54** 4562]
[12] Wang X M, Zhang J S, Zhang W F 2003 *Acta Phys. Sin.* **52** 2737 (in Chinese) [王小敏、张家树、张文方 2003 物理学报 **52** 2737]
[13] Zhan H, Wang X F, Li C H *et al* 2005 *Acta Phys. Sin.* **54** 4006 (in Chinese) [张 瀚、王秀峰、李朝晖等 2005 物理学报 **54** 4006]
[14] Li S J, Mou X Q 2003 *Comput. Phys. Commun.* **153** 52
[15] Li S J, Chen G R 2004 *IEEE Trans. Circ. Syst.* Ⅱ **51** 665
[16] Xun Y, Chik H T, Chee K S 2002 *IEEE Trans. Circ. Syst.* Ⅱ **49** 1826
[17] González J A, Reyes L I, Suárez J J *et al* 2002 *Physica* A **316** 259
[18] Zhou H, Ling X T 1997 *Acta Electron. Sin.* **25** 95 (in Chinese) [周 红、凌燮亭 1997 电子学报 **25** 95]
[19] Kawamoto S, Horiuchi T 2004 *Int. J. Bifurc. Chaos* **14** 3607
[20] Mao W B 2004 *Modern Cryptography: Theory and Practice* (Beijing: Electronic Industry Press) p108 (in Chinese) [毛文波 2004 现代密码学理论与实践 (北京:电子工业出版社) 第 108 页]
[21] Lasota A, Mackey M C 1994 *Chaos Fractal and Noise: Stochastic Aspects at Dynamics* (New York: Springer-Verlag)
[22] Knuth D E 2002 *The Art of Computer Programming* (New Jersey: Prentice Hall)

A novel chaos-based pseudo-random number generator*

Wang Lei Wang Fu-Ping Wang Zan-Ji
(*State Key Laboratory of Power System, Department of Electrical Engineering and Applied Electronic Technology, Tsinghua University, Beijing* 100084, *China*)
(Received 23 November 2005; revised manuscript received 10 April 2006)

Abstract

A novel pseudo-random number generator based on z-logistic map is proposed. The observed binary sequence of the chaotic orbit which is realized exactly under finite computing precision mostly retains the statistical characteristics and the randomness of the chaos-based information source which is defined on real domain, so the cryptographic properties of this novel chaos-based pseudo-random number generator (CPRNG) can be supported theoretically. Moreover, the period of this CPRNG is predicable and the weak keys can be excluded using a simple algorithm. This CPRNG overcomes the disadvantage of the traditional CPRNG whose existing weak keys are difficult to be excluded. Theoretical analysis and simulation results demonstrate that the cryptographic properties of the novel CPRNG are good, so it has potential application prospect in many areas including data encryption.

Keywords: chaos, pseudo-random number generator, information source
PACC: 0545

* Project supported by the Major Program of Science and Technology Research of Ministry of Education of China (Grant No. 105004).