基于人工免疫的网络入侵检测中疫苗算子的作用研究

方贤进[1,2] 李龙澍[3] 钱 海[4]
(安徽理工大学计算机科学与工程学院 淮南 232001)[1]
(安徽大学智能计算与信号处理教育部重点实验室 合肥 230039)[2]
(安徽大学计算机科学与技术学院 合肥 230039)[3]
(中国科学技术大学计算机科学与技术系 合肥 230039)[4]

摘 要 在以前研究工作的基础上,将包含疫苗算子、变异算子和其它算子的免疫算法与人工免疫中的负选择算法结合在一起,实现检测器种群的进化,目的是加快种群的亲和力成熟进程和提高网络入侵检测效率。详细地给出了疫苗自适应提取算法和疫苗算子算法,建立了基于免疫算法和负选择算法的模型及算法来实现网络入侵检测。分别设计了基于克隆选择算法的和基于免疫算法的网络入侵检测实验。实验结果表明,含有免疫算子的免疫算法加快了检测器种群亲和力成熟的进程,收敛速度更快,随着进化代数的增加检测率总体呈上升趋势。而基于克隆选择算法的网络入侵检测则出现了检测器种群亲和力成熟进程较慢,并随着进化代数的增加检测率呈现轻微退化和较长时间停滞不前的现象。

关键词 免疫算法,疫苗算子,克隆选择算法,网络入侵检测
中图法分类号 TP391 **文献标识码** A

Investigating the Role of Vaccine Operator in Artificial Immune System for Network Intrusion Detection

FANG Xian-jin[1,2] LI Long-shu[3] QIAN Hai[4]
(School of Computer Science and Engineering, Anhui Univ. of Science and Technology, Huainan 232001, China)[1]
(Key Lab of Intelligent Computing & Signal Processing of Education Ministry of Anhui Univ., Hefei 230039, China)[2]
(School of Computer Science & Technology, Anhui Univ., Hefei 230039, China)[3]
(Department of Computer Science & Technology of USTC, Hefei 230039, China)[4]

Abstract On the basis of author's previous research works, this paper integrates Immune Algorithm (IA) consisting of vaccine operator, mutation operator and other operators with Negative Selection Algorithm (NSA), called IA-NSA algorithm, to perform evolution of detector population, whose purposes are to quicken the process of affinity maturation and improve the efficiency of network intrusion detection. The algorithms of vaccine operator and adaptive extracting vaccines were given in detail. A novel network intrusion detection model and its algorithm based on IA-NSA were presented to establish anomaly detection. The network intrusion detection experiments based on IA-NSA algorithm and Clonal Selection Algorithm (CSA) were designed respectively, the experimental results show that the IA-NSA algorithm consisting of vaccine operator quickens the affinity maturation process of detector population and stably increases the detection rate along with increasing evolutionary generation. But in CSA-NSA algorithm, the affinity maturation process of detector population takes more time, the detection rate has a little degradation and maintains invariant for a long time.

Keywords Immune algorithm, Vaccine operator, Clonal selection algorithm, Network intrusion detection

当前许多计算机科学家提出了计算机免疫模型来解决各种应用问题,包括故障诊断、病毒检测和欺诈检测。其中,基于人工免疫系统的入侵检测的研究是最富有活力的[1,2]。

Hofmeyr提出的LISYS系统包含了利用负选择算法来生成初始检测器(immature detectors)种群。为了实现其自适应性,还提出了成熟检测器(mature detectors)、记忆检测器(memory detectors)、免疫耐受期(Tolerization period)、激活阈值(Activation threshold)和生命期(Life span)等机制[3]。LISYS模型中没有提出检测器种群的进化思想。

Kim提出了一个完整的网络入侵检测的人工免疫模型,该模型包括3个不同的进化阶段:负选择(Negative Selection Algorithm, NSA)、克隆选择(Clonal Selection Algorithm, CSA)和基因库进化(gene library evolution)。但该模型中的负选择不同于LISYS系统的,前者是利用NSA来过滤通过

到稿日期:2009-02-27 返修日期:2009-05-02 本文受安徽省科技攻关计划重大科技专项项目(08010201002),安徽省高等学校自然科学基金研究项目(kj2007B242)资助。
方贤进(1970—),男,博士研究生,副教授,主要从事智能计算与信息安全方面的研究,E-mail:xjfang@aust.edu.cn;李龙澍(1957—),男,教授,博士生导师,主要从事智能软件方面的研究;钱 海(1972—),男,博士,主要从事机器学习方面的研究。

基因表达和基因重组形成的预检测器(Pre-detectors);后者主要依靠 NSA 来生成不成熟的检测器群体。

在基于人工免疫系统(Artificial Immune System,AIS)的网络入侵检测领域,Kim 的贡献在于提出了检测器种群的进化思想——使用克隆选择算法(CSA)[4,5]。但在为检测器种群提供进化机会的同时,也不可避免地产生了退化现象,例如检测器多样性下降、抗体(检测器)亲和力下降和检测率(True Positive,TP)停滞不前等。另一方面,每个实际应用领域都会有一些基本的显而易见的特征信息或知识,尤其在网络入侵检测领域,一些先验知识对入侵检测起着很大的辅助和启示作用,忽视这些特征信息或知识所带来的损失是比较明显的[6]。

因此文献[7]提出了一个新的基于 AIS 的网络入侵检测模型,该模型集成了 NSA,CSA 与接种疫苗算子(Vaccine operator),具有分布式、自组织和轻量级的特性。本文就是在文献[7]的基础上研究由疫苗算子、变异算子及其它算子组成的免疫算法(Immune Algorithm,IA)在基于人工免疫的网络入侵检测中的作用,包括对检测器群体的亲和力成熟进程、检测率(TP)以及虚警率(False Positive,FP)的影响,并通过实验与基于 CSA-NSA 算法的 Kim 的网络入侵检测模型进行比较。

1 疫苗算子

疫苗算子是由接种疫苗和免疫选择两部分操作构成的。其中疫苗(vaccine)是指对所求解问题进行具体分析后从中提取出的最基本的特征信息。疫苗的正确选择对求解问题的效率具有十分重要的意义。

本文定义疫苗 $v=g_1,g_2,\cdots,g_m$,其中 $g_1,g_2,\cdots,g_m$ 是组成疫苗的基因,m 为组成疫苗的基因长度。疫苗可以根据先验知识构造或者用疫苗自适应提取算法生成,参见本文第 2 节。

1.1 接种疫苗

设检测器群体 $D=\{D^i \mid i=1,2,\cdots,n, D^i=d_1,d_2,\cdots,d_m\}$,其中 n 为检测器种群的大小,D^i 为检测器个体,$d_1,d_2,\cdots,d_m$ 为组成个体的基因,m 为组成检测器个体的基因长度。

假设 D_k^i 为第 k 代检测器群体 D 的第 i 个检测器个体,P_v 为检测器个体的接种概率,$V(D_k^i,g_j)$ 表示按模式 g_j 修改 D_k^i 上的第 j 个基因。

定义 1(接种疫苗) 指按某一概率 P_v 修改 D_k^i 的某些基因位(locus)上的基因,使所得到的个体以较大的概率具有更高的亲和力。对检测器个体 D_k^i 接种疫苗 v 的操作可表示如下:

```
vaccination(D_k^i, v):
    if Pv=true {
    j=random(m); D_k^i=V(D_k^i, g_j);}
```

random(m)表示产生 1,…,m 之间的任意一个正整数,其值表示接种疫苗的位置。

对 k 代检测器种群 D_k 进行接种疫苗是按比例 β 随机地抽取 $n_0=n\beta$ 个个体接种疫苗 v,形成中间种群 D_k'。该过程可表示为:

```
For i=1 to nβ {
    Randomly select a detector D_k^i from D_k;
    D_k'=D_k'+{Vaccination(D_k^i, v) };}
```

1.2 免疫选择

免疫选择包括两个步骤:第一步是免疫检查,第二步是免疫选择。免疫检查是对接种了疫苗的个体进行亲和力评估。如果该个体接种了疫苗后的亲和力不如父代,则该个体被其父代种群中相应个体取代,否则进入临时种群。免疫检查的过程可表示如下:

```
Immune inspection ( )
D_k^i'=vaccination(D_k^i, v);
//f is function of evaluating affinity
If f(D_k^i')<f(D_k^i);
  D_k^i'=D_k^i;
D_k'=D_k'+{D_k^i'};//enter into temporary population
```

免疫选择 immune selection( )是根据某种选择算子,如退火选择或杰出者选择,使用某种选择策略,如父子混合选择策略,从临时种群中选择一些个体进入下一代种群。

2 疫苗的生成

在入侵检测领域中,疫苗可以根据已知的攻击,从中提取表示攻击的特征信息生成,并对特征值进行编码构造疫苗。也可以在入侵检测的过程中动态地、自适应地提取疫苗。假设 $D_k=\{D_k^i \mid i=1,\cdots,n\}$ 是 k 代检测器种群,n 是种群大小。提取出的疫苗用 v 表示,$D_k^{optimal}$ 表示第 k 代种群中的当前最佳个体,m 表示组成疫苗的基因长度,f 是亲和力度量函数。则疫苗自适应提取算法用下面的伪代码表示:

```
extracting-vaccine ( )
for i=1 to n //find the optimal individual
  If f(D_k^i) > f(D_k^optimal) D_k^optimal=D_k^i;
//decompose the optimal individual and get vaccine
for j=1 to m do v=v+D_k^optimal(j);
return v
```

3 对 k 代检测器种群执行疫苗算子

根据本文中第 1 节和第 2 节,疫苗算子可以用如下的随机过程来描述:

$$D_k \xrightarrow[\text{with ration } \beta]{\text{selection}} D_k(n\beta) \xrightarrow[\text{immune inspection}]{\text{vaccination}} D_k'(n\beta) \cup D_k \xrightarrow{\text{immune selection}} D_{k+1}$$

对 k 代检测器种群执行疫苗算子的算法可用图 1 中的代码表示。

```
Step1  v=extracting-vaccine();
Step2  randomly selection nβ detectors from D_k with ratio β, get temporary population D_k(nβ);
Step3  for j=1 to nβ {
         select a detector D_k^i from D_k(nβ);
         D_k^i'=vaccination(D_k^i, v);
         immune inspection();
         D_k'(nβ)=D_k'(nβ)+{D_k^i'};}
Step4  D_k''=D_k'(nβ) ∪ D_k;
       //perform selection operator
Step5  D_{k+1}=selection(D_k'').
```

图 1 对 k 代检测器种群的疫苗算子操作

4 对检测器种群的免疫算法

对检测器种群执行的免疫算法(IA)包括变异算子、疫苗算子和选择算子,算法如图 2 所示。

Step1 create initial detector population D by gene expression or recombination[4];

Step2 if satisfying termination criterion (e. g. requirement of detection rate) or maximum evolutionary generation, then algorithm ends, otherwise continues;

Step3 v=extracting-vaccine( );

Step4 D_k'=mutation(D_k);

Step5 D_k''=vaccine_operator(D_k',v);

Step6 D_{k+1}=Selection(D_k'')

Step7 perform next network intrusion detection;

Step8 go to step2.

图 2 对检测器种群执行的免疫算法(IA)

在图 2 的免疫算法中,没有使用交叉算子,这是因为在免疫系统中抗体的亲和力成熟主要是通过超级变异(hypermutation)来实现的。

5 免疫算法和负选择算法的集成实现网络异常检测

在检测器群体进化过程中生成的检测器有可能与自身轮廓(self profile)相匹配,所以要使用 NSA 对检测器个体进行过滤,删除那些与自身匹配的检测器,这样可以降低虚警率(FP)。

将 IA 和 AIS 中的 NSA 集成,实现网络入侵检测的算法如图 3 所示,这个算法被称为 IA-NSA 算法。此算法中 NSA 同样只是作为子代检测器的过滤器,而不是用来随机生成初始检测器群体。这与 Kim 的基于人工免疫的网络入侵检测模型是一样的。

Step1 create SELF Profile (it represents normal network traffics);

Step2 create initial detector population D;

Step3 if satisfying termination criterion (e. g. requirement of detection rate) or maximum evolutionary generation, then algorithm ends, otherwise continues;

Step4 v=extracting-vaccine( );

Step5 D_k'=mutation(D_k);

Step6 D_k''=vaccine_operator(D'_k,v);

Step7 D_{k+1}=*Selection*(D_k'');

Step8 perform negative selection algorithm for population D_{k+1} under a certain match rule r.

for i=1 to n

for each $s\in SELF$ do

If $match_r(D_{k+1},s)$=true remove D_{k+1} from D_{k+1};

Step9 perform next network intrusion detection;

Step10 go to step3.

图 3 免疫算法与负选择算法的集成(IA-NSA 算法)

6 仿真实验

6.1 实验数据

实验数据来自 KDDCUP1999 入侵检测数据集[8],训练集取 10%子集共 494021 个样本,含有 4 大类(DoS, Probe, U2R, R2L)共 22 种类型的攻击,攻击样本 396743 个,正常网络数据样本 97278 个。测试集是带标注攻击类型的 10%子集,共 311029 个样本,其中含有 37 种攻击类型(有 15 种攻击类型未在训练集中出现)共 250436 个样本,另外含有正常网络数据 60593 个样本。

6.2 实验目标

本文中的免疫算法 IA 与负选择算法 NSA 的结合称为 IA-NSA 算法,Kim 模型[4] 中的 CSA 与 NSA 的结合称为 CSA-NSA 算法。本实验目标就是研究包含疫苗算子的 IA-NSA 算法在基于人工免疫的网络入侵检测中的作用,包括对检测器群体的亲和力成熟、检测率、虚警率的影响,并通过实验结果与 CSA-NSA 算法进行比较。

6.3 实验设计及参数设置

6.3.1 初始检测器种群的生成

考虑到负选择算法的效率,初始检测器种群的生成没有采用随机生成的方法,而是根据文献[4]提出的思想,利用训练集中已知攻击数据的特征,抽取基因建立基因库,再利用基因表达和基因重组生成初始检测器种群,种群大小为 200。

6.3.2 检测器个体编码

KDDCUP1999 入侵检测数据集中每个样本含有 41 个特征(feature),有些特征值是连续数据,有些是离散数据。为进行进化计算,本文采用二进制编码,因而需要对初始检测器种群中的连续特征值进行离散化。离散化算法采用 Naïve Scaler 算法[9]。每个检测器个体含有 41 个基因,检测器个体编码长度 L=104bits。

6.3.3 匹配函数

检测器个体的亲和力计算以及对网络数据进行异常检测都要用到匹配函数。本文采用改进的 r-contiguous bits 匹配规则[10]作为匹配函数:

$$match_r(Ab,Ag)=\begin{cases} success, & \exists k,1\leqslant k\leqslant l-r+1,\bigcap_{m=k}^{k+r-1}Ab.g_m=Ag.g_m \\ failure, & \text{otherwise} \end{cases}$$

其中,Ab 表示抗体(检测器),Ag 表示抗原,g_m 表示第 m 个基因,l 表示组成抗体的基因个数;r 是匹配阈值,但它不再表示连续匹配的比特数,而是表示连续匹配的特征值(即基因的表现型)的个数。本文中设置匹配阈值 r=12。

6.3.4 检测器个体亲和力计算方法

本文中采用如下算法进行个体亲和力计算。首先从异常样本集中随机抽取 *asize* 个组成抗原集 A。然后对检测器种群中的每个个体 D^i 与 A 中的所有抗原 Ag 进行匹配并计数。最后用匹配成功的数量占总的抗原数量的比例作为该检测器个体 D^i 的亲和力。

```
for i=1 to asize {
Randomly select Ag_i from abnormal;
A=A+{ Ag_i };}
for each D^i in detector population D {
  match_count=0;
  For i=1 to asize {
  If match_r(D^i,Ag_i)=true
  match_count=match_count+1;}
  D^i. affinity=match_count/asize;
  }
```

其中，$D.affinity$ 表示检测器个体的亲和力，$match_r(\ )$ 是在某一匹配规则 r 下的匹配函数，$match\_count$ 是用来计数的变量。实验中设置的非自身抗原样本集大小为 $asize=200$。

6.3.5 变异算子、选择算子、进化代数和接种疫苗比例

变异采用均匀点变异，变异概率 $P_m=0.6$。选择算子采用杰出者选择，并且采用父子混合选择策略[11]，这样可提高检测器种群的多样性。进化代数 $G=100$，接种疫苗比例 $\beta=0.8$。

6.3.6 克隆选择算法中的有关参数设置

为了将 IA-NSA 算法与 CSA-NSA 算法在网络入侵检测方面的性能进行比较，需要对同一数据集进行克隆选择算法实验。本文使用静态 CSA[5]。有关参数设置如下：变异和选择算子设置同上，匹配阈值设置同上。在进化时，每次从种群中选择 20%的最优个体组成最优种群，然后执行 CSA 算法，在新生成的检测器中选择 20%的最优个体代替检测器种群中亲和力低的旧检测器。克隆规模与检测器的亲和力成正比，克隆比例 $Mc=5$。

6.4 实验结果及分析

6.4.1 两种算法 IA-NSA 和 CSA-NSA 对检测器种群亲和力的影响

实验结果 1 是在种群大小 $POPSize=200$，变异概率 $P_m=0.6$，接种疫苗概率 $\beta=0.8$，匹配阈值 $r=12$，克隆系数 $Mc=5$，选择最优种群比例 $bestM=0.2$ 的情况下，对同样的初始检测器种群分别执行 IA-NSA 算法和 CSA-NSA 算法（进化代数 $G=100$），种群的亲和力成熟过程及分布情况如图 4 和图 5 所示。

通过实验结果统计出，在 IA-NSA 算法中，当检测器种群进化到第 50 代时，检测器种群中有 45 个检测器个体的亲和力收敛到 1。而当进化到 100 代时，种群中有 85 个检测器个体的亲和力收敛到 1。在 CSA-NSA 算法中，当检测器种群进化到 50 代时，检测器个体的亲和力最高为 0.90；而当进化到 100 代时，检测器个体的亲和力最高为 0.98。可见，两种算法对检测器亲和力成熟进程的影响是不一样的，IA-NSA 算法要更快一些。

在 CSA-NSA 算法的迭代搜索过程中，尽管使用了基于亲和力大小的比例选择来产生最优解子种群，但其变异仍然是在一定的概率条件下随机地、无指导地实现，具有一定的盲目性，做了一些无效的工作。另外，CSA-NSA 算法中由于使用了克隆算子，可能会产生重复或拥挤解，也影响了群体的亲和力成熟进程。

IA-NSA 算法由于使用了接种疫苗算子，迭代搜索过程有指导地进行，使得个体以较大的概率具有更高的亲和力，群体的亲和力也相对稳定地提高，其抑制或避免了亲和力成熟过程中的一些重复或无效的工作，提高了亲和力成熟的效率。

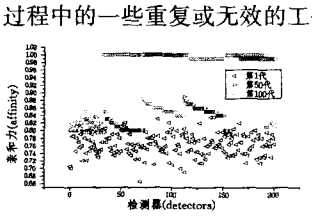

图 4 IA-NSA 算法对检测器种群亲和力成熟过程的影响

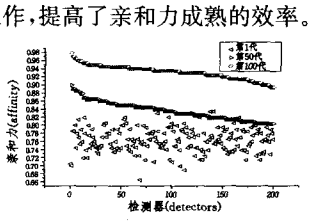

图 5 CSA-NSA 算法对检测器种群亲和力成熟过程的影响

6.4.2 IA-NSA 和 CSA-NSA 算法的检测率和虚警率

使用 IA-NSA 算法对入侵检测测试集进行检测所得到的检测率（TP）和虚警率（FP）如图 6 所示。

使用 CSA-NSA 算法对入侵检测测试集进行检测所得到的检测率（TP）和虚警率（FP）如图 7 所示。

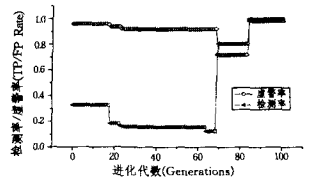

图 6 使用 IA-NSA 算法进行检测得到的 TP 和 FP

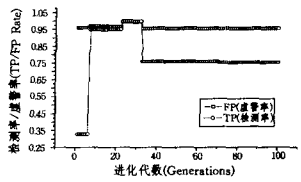

图 7 使用 CSA-NSA 算法进行检测得到的 TP 和 FP

首先从图 6 和图 7 可以看出，无论是 IA-NSA 算法还是 CSA-NSA 算法，进行入侵检测时均具有较高的虚警率（FP），这是因为本文中没有采用检测器的耐受期（Tolerization Period）、激活阈值（Activation Threshold）、生命期（Life Span）等其它机制，这也是下一步工作要解决的问题。

另外从图 6 可以看出，在检测器种群 1～100 代的进化过程中，IA-NSA 算法的检测率总体还是呈上升趋势，例如在第 69 代达到 80.51%，在第 84 代达到 99.76%，然后稳定在一个高检测率状态。图 7 中的 CSA-NSA 算法进行入侵检测时，当检测器群体进化到第 4 代时，检测率有很大的飞跃，达到 94.79%。但从第 33 代开始，检测率还是有了一定程度的退化，此后很长时间稳定在一个比较高的状态。

结束语 含有疫苗算子的 IA-NSA 算法与 CSA-NSA 算法分别进行网络入侵检测时，从抗原（检测器）亲和力成熟的角度来看，IA-NSA 算法具有收敛速度快、亲和力成熟进程快的明显优势。从入侵检测的检测率和虚警率角度来看，IA-NSA 算法随着进化代数的增加，其检测率呈总体上升趋势，而 CSA-NSA 算法在种群进化过程中，其检测率出现了较小的退化和较长时间停滞不前的现象。另外，两种算法尽管能达到较满意的检测率，但同时具有很高的虚警率。下一步应着重研究检测器的耐受期（Tolerization Period）、激活阈值（Activation Threshold）、生命期（Life Span）以及其它机制，并实现虚警率的降低。

参 考 文 献

[1] Dasgupta D. An Overview of Artificial Immune Systems and Their Applications[M]. Artificial Immune Systems and Their Applications, Berlin: Springer-Verlag, 1998: 3-21

[2] Somayaji A, Hofmeyr S, Forrest S. Principles of a Computer Immune System[C]// Proceedings of the 1997 Workshop on New Security. New York: The Association for Computing Machinery, 1997: 75-82

[3] Hofmeyr S A. An Immunological Model of Distributed Detection and Its Application to Network Security[D]. University of New Mexico, 1999

[4] Kim J. The artificial immune model for network intrusion detection[C]// Proc. of 7th European Congress on Intelligent Techniques and Soft Computing, Aachen, Germany, 1999

（下转第 270 页）

视频镜头的平均帧间差小于 T，则说明该镜头内容变化非常缓慢，为了计算方便就取镜头的中间帧作为关键帧，算法结束。若视频的平均帧间差大于 T，就进入流程(2)。

(2)初始化偏向参数 $p=S_{sort}(pmedia+b)$，定义最大循环次数 $mcount=50000$，收敛条件为聚类中心 50 次循环无变化，初始化吸引度 R_i 和归属度 A_i 为 0，每次偏向参数 p 下调步伐 $p_{step}=0.1*p_{min}$（p_{min} 是偏向参数 p 的最小值）。

(3)按照式(4)－式(8)进行循环迭代，不断更新吸引度 R_i 和归属度 A_i，算法产生 k 个聚类中心。

(4)判断是否收敛，若收敛或者是趋于收敛，则转入流程(5)，若聚类个数不下降且聚类指标值下降，则 $p=p+p_{step}$，转入流程(5)，否则转入流程(3)。

(5)计算 Silhouette 指标值，判断 k 是否小于等于 2，循环次数是否等于 $mcount$，若聚类中心个数小于等于 2，或者循环次数达到最大值，则转入流程(6)，否则减小参数 $p=p+p_{step}$，转入流程(3)。

(6)找出 sil 最大值所对应的聚类个数，即最优的聚类结果，其聚类中心就是最优的关键帧。

5 实验结果

为了验证上述算法的有效性，采用了不同类型的实验视频，包括：MTV、足球、广告、新闻不同类型的视频，视频序列从几百帧到几万帧，视频帧大小各不相同。其基本情况如表 1 所列。

表 1 实验数据集

样本	帧数	本算法检测出的关键帧数
MTV	449	34
足球	3525	85
广告	577	24
新闻	14862	71

以两段典型的视频为例，在视频内容变化缓慢的广告视频第一个镜头中，应用本算法找出的关键帧如图 1 所示。

图 1 提取的广告关键帧

镜头 1 有 168 帧，应用本算法提取的关键帧有 3 帧，反映了视频的运动信息，包括昂头、平视、侧视这 3 个动作，没有冗余。

在足球运动视频中，以其中一个目标运动变化剧烈的镜头为例，应用本算法找出的关键帧如图 2 所示。

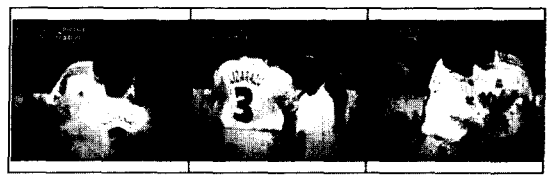

图 2 提取的足球运动关键帧

我们可以看到通过这 3 帧反映了人物的移动，表达了视频镜头的主要内容，尤其是运动信息。

从以上的实验结果分析可以看到，不管是对目标对象运动缓慢的视频，还是对目标运动剧烈、镜头快速移动的视频，本算法都能有效地找出代表视频主要内容和运动信息的最优关键帧，而且算法稳健快速。

结束语 本文提出了基于仿射传播聚类的自适应关键帧提取算法，克服了运用 K 均值算法时难以确定聚类个数以及初始聚类中心的随机性带来的结果的不稳定性等缺陷，本算法能够根据视频本身的信息分布以及视频的复杂度，自适应地提取出最优关键帧，在一系列不同类型的视频片断上进行实验和分析，实验的提取结果与人的视觉有良好的一致性，证明了该算法能快速有效地提取出关键帧，并且具有良好的鲁棒性。

参 考 文 献

[1] Zhang Hongjiang, Wang J Y A, Altunbasak Y. Content - based video retrieval and compression: A unifed solution[A]// IEEE International Conference on Image Processing[C]. Washington, DC, 1997: 13-16

[2] Worf W. Key frame selection by motion analysis[C]// Proceedings of the 1996 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Atlanta, 1996: 1228-1231

[3] 伯晓晨，刘建平. 基于颜色直方图的图象检索[J]. 中国图像图形学报，1999，4(1)：33-37

[4] 邢伟利. 图像检索中颜色的特征提取及匹配算法[J]. 微机发展，2002(2)：86-88

[5] Frey B J, Dueck D. Clustering by passing messages between data points[J]. Science, 2007, 315(5814): 972-976

[6] 王开军，张军英，李丹，等. 自适应仿射传播聚类[J]. 自动化学报，2007，12(33)：1242-1245

[7] Velamuru P K, Renaut R A, Guo H B, et al. Robust clustering of positron emission tomography data[C]// Joint Interface CSNA. USA, 2005

(上接第 242 页)

[5] Kim J. Towards an artificial immune system for network intrusion detection: an investigation of clonal selection with a negative selection operator[C]// Proc. of the Congress on Evolutionary Computation. Seoul, Korea, 2001: 1244-1252

[6] 焦李成，杜海峰，刘芳，等. 免疫优化计算、学习与识别[M]. 北京：科学出版社，2006

[7] Fang Xianjin. An Artificial Immune Model with Vaccine operator for Network Intrusion detection[C]// Proc. of 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application. IEEE CS Press, 2008: 488-491

[8] Hettich S, Bay S D. The UCI KDD Archive[OL]. Irvine, CA: University of California, Department of Information and Computer Science. http://kdd.ics.uci.edu/database/kddcup99/1999

[9] 施伟，战守义，盛思源. 基于 Rough Set 的数据预处理[J]. 计算机工程与应用，2003，39(22)：190-191

[10] Forrest S, Allen L, Cherukuri R. Self-nonself discrimination in a computer[C]// Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy. IEEE Computer Society Press, 1994: 271-281

[11] 徐宗本，张讲社，郑亚林. 计算智能中的仿生学：理论与算法[M]. 北京：科学出版社，2003