

Biological Immune System

- depending on type of pathogen and its way of entry, different response mechanisms used:
 - neutralize pathogenic effect
 - destroy infected cells
- immune system:
 - innate immune system: present at birth
 - adaptive immune system: learning,
 - adaptability and memory

Biological Immune System

- immune system features relevant to AIS
- matching: binding between antigen and antibody
- diversity: in order to achieve optimal antign space coverage, antibody diversity encouraged
- distributed control: immune system governed by local interactions between immune cells and antigens

Immune Network Theory

- immune system maintains idiotypic network of interconnected B-cells for antigen recognition
 - B-cells responsible for antibody (proteins that bind to antigens) production
- suppress and stimulate each other in such a way as to lead to stabilization of network

Immune Network Theory

- two B-cells are connected if the affinities they share exceed a threshold
- strength of the connection directly proportional to affinity they share

Negative Selection Mechanism

- aim is to provide tolerance for self-cells
- deals with immune system's ability to: - detect unknown antigens
 - not react to self-cells

Negative Selection Mechanism

- when immune system cells are created they undergo negative selection in thymus
- those that react against self-cells are destroyed
- those that do not are released from thymus

Clonal Selection

- describes basic features of an immune response to an antigen
- only those that recognize an antigen proliferate (increase in no.s)
 - selection against those that do not match

Clonal Selection

- main features of clonal selection theory:
 - new cells which are copies (clones) of parents are subject to a high mutation (somatic hypermutation)
 - elimination of newly differentiated lymphocytes carrying self-reactive receptors.
 - proliferation and differentiation of mature cells

Illustrative Problems

- intrusion detection
- data mining collaborative filtering and clustering

Intrusion Detection

- to identify potential attacks and to react by generating an alert or by blocking the possible attack
- main goal is to detect misuse, unauthorized use and abuse of computer systems by both insiders and external intruders

Intrusion Detection

- current IDS define suspicious signatures based on known intrusions and probes
 - limitation: failure to detect previously unknown attacks

Data Mining – Collaborative Filtering (CF) and Clustering

- broad range of algorithms that use a similarity measure to obtain recommendations
 - e.g. "people who bought this also bought these"
- any problem domain where users are required to rate items is suitable for CF
 – recommender systems
- related problem: clustering of similar preferences or types

Artificial Immune Systems

- decisions when implementing AIS:
 - encoding
 - similarity measure
 - selection
 - mutation

Encoding

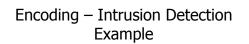
- in an application domain
 - antigen = target solution
 - data item to check if it is an intrusion
 the user which is being clustered or is being made a recommendation for
 - antibodies: remainder of data
 - other users in database
 - network traffic which has already been identified

Encoding

- there can be more than one antigen and usually many antibodies
- antigens and antibodies are repesented in the same way
- efficient encoding / representation is chosen similar to in EAs

Encoding – Data Mining Example

- movie recomender
- user profile should be represented
 - movies seen
 - how much liked
- possible encoding:
- $\{\{id1, score1\}, \{id2, score2\}, \ldots, \{idn, scoren\}\}$



• for each data packet transferred:

[<protocol><source ip><source port><destination ip><destination port>]

- example:
- [<tcp><113.112.255.254><any><108.100.111.12>
 <25>]

Similarity / Affinity Measure

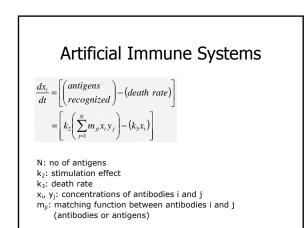
- similarity measure = matching rule
 - related to encoding
 - e.g. for binary encoding may use Hamming distance or the length of the longest of the continuous no. of matching bits
 - e.g. for real encoding, may use Euclidean distances

Similarity / Affinity Measure

- for data mining applications similarity usually means correlation
 - e.g. use Pearson correlation coefficient between users
 - -1 means strong disagreement, 1 means strong agreement, 0 means no correlation

Artificial Immune Systems

- 1. initially AIS is empty
- 2. target user is set as antigen, all others as antibodies
- 3. add antigen to AIS
- 4. add one candidate antibody at a time
- antibodies start with an initial concentration which decreases over time (death rate)
 - low concentration antibodies removed
 high concentration antibodies may go to catural
 - high concentration antibodies may go to saturation
 matching antibodies increase in amount (better matching ones increase more) ⇒ stimulation



Pseudocode of AIS for the Movie Recommender System

initialize AIS;

- encode user to make predictions for as Ag; while (AIS not full) and (more antibodies) do
- add next user as antibody Ab; calculate match score between Ag and Ab; while (AIS at full size) and (AIS not stabilized) do
- { reduce concentration of all Abs by a fixed amount;
- match each Ab against Ag; stimulate as necessary;
 }
- use final set of antibodies for recommendations;

Result of AIS for the Movie Recommender

- in the final set of antibodies, use concentrations as weights and take a weighted average to determine recommendation
- AIS is called stabilized if (for e.g.) for 10 iterations no size change occurs
- stabilization means that sufficient no of good neighbors have been found

 a good prediction can be made using it

Somatic Hypermutation

- similar to EAs

 but mutation is less distruptive for better matches
- mutation may not make sense for some types of data
 - mutating user to be more similar to target works but basing a recommendation on the artificial users is not meaningful

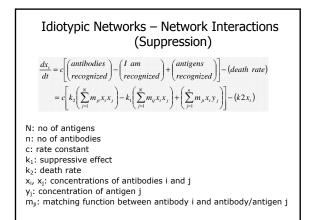
Somatic Hypermutation

- mutation may make sense for some types of data
 - meaningful in IDS: for example mutating those that match self rather than dscarding them directly
 - in clustering, cluster centers can be artificial users, so mutating cluster centers is meaningful

Idiotypic Networks – Network Interactions (Suppression)

- antibodies can match other antibodies as well as antigens
- antibody matches others which also match others too

 can explain memory of past infections
- this could result in suppression of similar antibodies to increase diversity



Pseudocode of AIS for the Movie Recommender System with Idiotypic Networks

initialize AIS; encode user to make predictions for as Ag; while (AIS not full) and (more antibodies) do { add next user as antibody Ab; calculate match score between (Ag and Ab) and (Ab and other Abs); while (AIS at full size) and (AIS not stabilized) do { reduce concentration of all Abs by a fixed amount; match each Ab against Ag; stimulate as necessary; match each Ab against each other and execute idiotypic effect; } } use final set of antibodies for recommendations;

Standard Negative Selection Algorithm

- 1. initialize random detectors;
- 2. censoring:
 - 1. match evaluation: determine affinity between every self and detector;
 - 2. selection: if a detector matches a self then discard it else add to detector set;
- 3. monitoring: monitor neew set of self for possible changes / variations;

Standard Clonal Selection Algorithm

- 1. initialize random detectors;
- 2. for each antigen do:
 - selection: select antibodies which have a higher affinity to antigen;
 - 2. reproduction and genetic variation:
 - reproduction and genetic variation.
 produce copies (clones) of selected cells(no. depends on how good the match is);
 mutate each inversely proportional to its affinity to antigen;
 - affinity evaluation;
- cycle: repeat step 2 until stopping criteria are met;