

A4. Совместимость фильтров Блума

Демченко Георгий Павлович , БПИ-235

1. $F(AB)$

Будем считать что для элементов из множеств A и B использовались одинаковые хеш-функции при составлении соответствующих фильтров Блума.

При таком подходе фильтр $F(AB)$ будет выдавать положительные ответы о принадлежности объектов из множества $A \cap B$, так как :

$$\forall c \in A \cap B : \forall h_i : (T_{F(A)}[h_i(c)] = T_{F(B)}[h_i(c)] = 1) \Rightarrow T_{F(AB)}[h_i(c)] = 1 \& 1 = 1$$

где

h_i - i -ая хэш-функция, использовавшаяся для вычисления хэша элемента при вставке в фильтр

T_F - внутренняя таблица соответствующего фильтра

$T_F[\dots]$ - элемент внутренней таблицы соответствующего фильтра

Говоря словами, если элемент $c \in A \cap B$, то соответствующие его хэсам элементы в фильтрах $F(A)$ и $F(B)$ будут иметь одинаковые позиции (индексы) в внутренних таблицах и будут равны 1. Тогда при образовании фильтра $F(AB)$ с помощью поэлементного побитового И они также перенесутся во внутреннюю таблицу $F(AB)$ на те же позиции (индесы), так как $1 \& 1 = 1$

Получаем

$$\forall c \in A \cap B : \forall h_i : T_{F(AB)}[h_i(c)] = 1 \Rightarrow F(AB)[c] = true$$

Ответ: Да, верно

2. $F(AB) \neq F(A \cap B)$

Нет, $F(AB)$ не всегда будет в точности соответствовать фильтру $F(A \cap B)$, полученному вставкой объектов из множества $A \cap B$, так как :

Таблица фильтра $F(AB)$ может содержать лишние биты, по сравнению с таблицей $F(A \cap B)$, образованные пересечением 1 из таблиц фильтров $F(A)$ и $F(B)$ на позициях хешей элементов, которые не лежат в $A \cap B$.

Может произойти ситуация, где:

$$\exists a_1 \in A \setminus B, b_1 \in B \setminus A, a_1 \neq b_1 : \exists h_n : h_n(a_1) = h_n(b_1)$$

$$\Rightarrow T_{F(A)}[h_n(a_1)] = T_{F(B)}[h_n(b_1)] = 1$$

$$\Rightarrow T_{F(AB)}[h_n(b_1)] = 1 \ \& \ 1 = 1$$

Т.е. образуется коллизия элементов, не принадлежащих $A \cap B$, в таблицах $F(A)$ и $F(B)$, тогда в таблице $F(AB)$ появятся лишние 1 на позициях, которые не соответствуют позициям из $F(A \cap B)$, что может дополнительно (помимо коллизий) привести к ложноположительному ответу $F(AB)$ на принадлежность элементу не из $A \cap B$ множеству $A \cap B$.

Ответ: Нет, неверно