



Підхід до моделювання процесів вирішення тематичних задач ДЗЗ на основі теорії свідчень

М.О. Попов (1), О.В. Зайцев (2)

(1) Науковий центр аерокосмічних досліджень Землі ІГН НАН України

(2) Воєнно-дипломатична академія імені Євгенія Березняка

Четверта міжнародна конференція “Аерокосмічні спостереження в інтересах сталого розвитку та безпеки”, Київ, 26–30 травня 2014 р.



БАЗОВІ СКЛАДОВІ ЗАБЕЗПЕЧЕННЯ СТАЛОГО РОЗВИТКУ



Економічна складова – оптимальне використання природних ресурсів і використання екологічних технологій, включаючи добування та переробку сировини, створення екологічно прийнятної продукції, мінімізацію, переробку та знищення відходів.

Соціальна складова – збереження стабільності соціальних і культурних систем, в тому числі, скорочення чисельності конфліктів.

Екологічна складова – забезпечує цілісність біологічних та фізичних природних систем.



ПРИРОДНІ РЕСУРСИ І ФАКТОРИ, ЩО ВИЗНАЧАЮТЬ СТАЛІЙ РОЗВИТОК



і можуть бути оцінені дистанційними методами:

- Геологія
- Землекористування
- Сільське господарство
- Екосистеми та лісогосподарство
- Океан і прибережні зони
- Водні ресурси
- Метеорологія та кліматологія
- Надзвичайні ситуації
- Національна безпека і оборона

Четверта міжнародна конференція “Аерокосмічні спостереження в інтересах сталого розвитку та безпеки”, Київ, 26–30 травня 2014 р.



СТРУКТУРА ЦИКЛУ ДИСТАНЦІЙНОГО ЗОНДУВАННЯ ЗЕМЛІ



Постановка завдання



Планування ДЗЗ



Реалізація ДЗЗ:

- Отримання даних ДЗЗ
- Первинна обробка даних



Аналітична обробка



Інтеграція даних



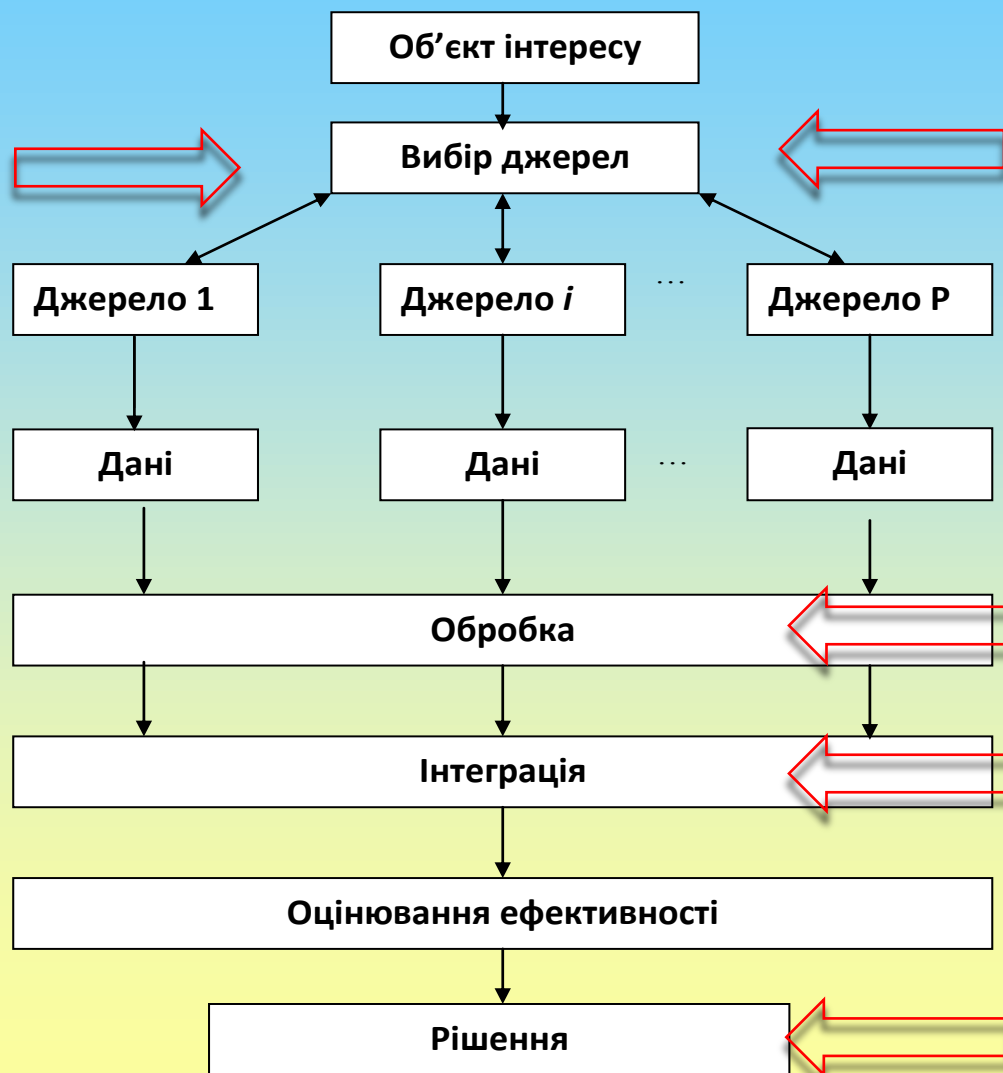
Прийняття рішення



ЗАГАЛЬНА СХЕМА ВИРІШЕННЯ ТЕМАТИЧНИХ ЗАДАЧ ДЗЗ



- Космічні знімки високого та надвисокого просторового розрізнення, багатоспектральні та гіперспектральні знімки (WorldView 1,2, GeoEye, Ikonos, QuickBird ...).
- Електронні карти, цифрові моделі рельєфу, композити попередніх знімків та ін.



- Обмежені можливості кожного з відомих засобів.
- Великі обсяги даних ДЗЗ та обмежений ресурс часу на їх оброблення та доставку користувачам.
- Вимоги щодо уніфікації форм подання та різнотипні дані.
- Зростаючі вимоги до достовірності та повноти інформації ДЗЗ.



НАУКОВО-МЕТОДИЧНИЙ АПАРАТ ДЛЯ ВИБОРУ, ІНТЕГРАЦІЇ ТА ПРИЙНЯТТЯ РІШЕННЯ



1. **Теорія ймовірностей** не дозволяє вирішити наукову задачу з наступних причин:
 - події та факти про об'єкти інтересу не відповідають вимогам до обсягів статистичних даних та вибірок;
 - статистичні методи не дозволяють обробляти незнання;
 - труднощі з обробленням даних від корельованих джерел.
2. **Теорія нечітких множин**: наявність суб'єктивного фактору при використанні функції належності, що потребує експертних оцінок.
3. **Теорія свідчень Демпстера-Шейфера** є узагальненням статистичної теорії і позбавлена зазначених вище недоліків.



КОМБІНУВАННЯ І ІНТЕГРАЦІЯ ДЖЕРЕЛ (правило комбінування Демпстера)



$$m_{DS}(X) = \frac{1}{1 - k_{12}} \cdot \sum_{\substack{X_1, X_2 \in 2^\Omega \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) \quad (1)$$

де:

$m_{DS}(X)$ - комбінована базова ймовірність 2-ох джерел

Ω - універсальна множина (основа аналізу)

$m_1(\cdot)$ - базова ймовірність фокального елементу 1-ого джерела

$m_2(\cdot)$ - базова ймовірність фокального елементу 2-ого джерела

$m_{DS}(\emptyset) = 0$, $\forall (X \neq \emptyset) \in 2^\Omega$ - умови

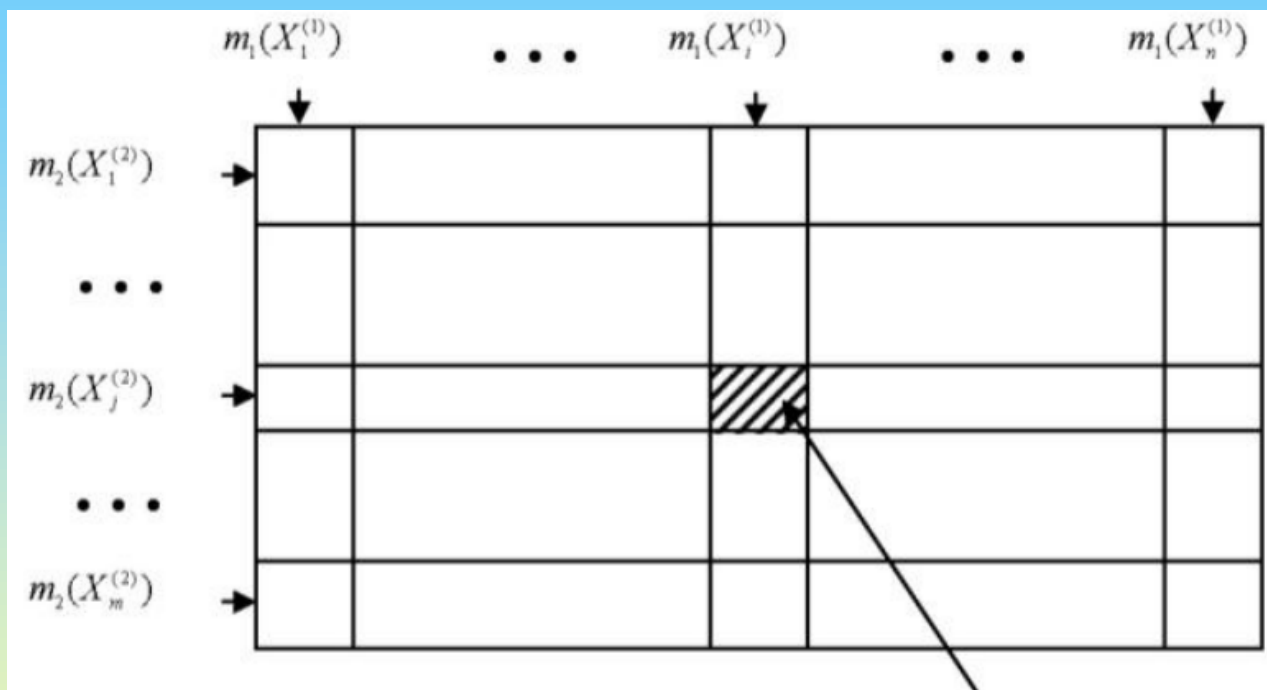
X_1, X_2 - групи свідчень, що отримані від 1-ого та 2-ого джерел

k_{12} - ступінь конфліктності

$$k_{12} = \sum_{\substack{X_1, X_2 \in 2^\Omega \\ X_1 \cap X_2 = \emptyset}} m_1(X_1) m_2(X_2) \quad (2)$$



НЕДОЛІКИ ПРАВИЛА КОМБІНУВАННЯ ДЕМПСТЕРА



Ортогональна сума ймовірностей двох джерел - $X_i^{(1)} \cap X_j^{(2)}$

- Якщо всі свідчення джерел мають протиріччя, тоді використання правила комбінування Демпстера не є коректним, так як ці джерела не можуть бути поєднані.
- Не враховується інформація, що отримана з конфліктуючих та ненадійних джерел.

Четверта міжнародна конференція "Аерокосмічні спостереження в інтересах сталого розвитку та безпеки", Київ, 26–30 травня 2014 р.



АЛЬТЕРНАТИВНІ ПРАВИЛА КОМБІНУВАННЯ ДЖЕРЕЛ



- Шейфера та Сметса: для врахування невизначеності при призначенні основних мас ймовірностей (на основі недостовірних свідчень з ненадійних джерел) пропонується їх дисконтування коефіцієнтом надійності.
- Ягера та Інагакі: правила використовують комбіновані маси ймовірностей по перетинам виділених підмножин для відображення ступеню незнання.
- Жанга: правило центрального комбінування, враховує ступінь перетину виділених підмножин, визначених на основі різних груп свідчень та вводить оцінку ступеню їх перетину.
- Дюбуа та Прада: діз'юнктивне правило комбінування. Має переваги коли відомо, що один або декілька джерел надають неправдиву інформацію, при цьому невідомо яке саме.
- Правила засновані на усередненні основних значень ймовірностей.



ХАРАКТЕРИСТИКИ ДЖЕРЕЛ ДАНИХ



Відомі правила комбінування, не враховують знання (метадані) про джерела інформації:

- Достовірність (*truthfulness*) – ступінь вірності даних, що надає джерело,
- Релевантність (*relevance*) – відповідність джерела поставленому завданню,
- Надійність (*reliability*) – характеристика адекватності джерела
- Завершеність (*completeness*) – ступінь компетентності джерела
- Чесність (*sincerity*) – ступінь того, наскільки джерело не приховує інформацію від споживача.



ВРАХУВАННЯ МЕТАДАНИХ ДЖЕРЕЛА



$$R = \{R, \neg R\}$$

– параметр джерела – релевантність;

$$T = \{T, \neg T\}$$

– параметр джерела – достовірність;

$$H = R \times T$$

– множина можливих станів джерела;

$$prob(h_R, h_T)$$

– мета-знання про джерело, можуть приймати форму суб'єктивних ймовірностей про стан джерела;

$$\Gamma_A(R, T) = A$$

$$\Gamma_A(R, \neg T) = \bar{A}$$

$$\Gamma_A(\neg R, T) = \Gamma(\neg R, \neg T) = X$$

– функція інтерпретує свідчення для кожного стану джерела h ;

(3)

$$m^X(A) = prob(R, T),$$

$$m^X(\bar{A}) = prob(R, \neg T),$$

$$m^X(X) = prob(\neg R) = prob(\neg R, T) + prob(\neg R, \neg T)$$

(4)



ПЕРЕТВОРЕННЯ ФУНКЦІЇ МАСИ З ВРАХУВАННЯМ МЕТАДАНИХ



$$\left\{ \begin{array}{l} \text{prob}(R, T) = pq \\ \text{prob}(R, \neg T) = p(1 - q) \\ \text{prob}(\neg R, T) = (1 - p)q \\ \text{prob}(\neg R, \neg T) = (1 - p)(1 - q) \end{array} \right. \quad (5)$$

$$\left\{ \begin{array}{l} m^X(A) = pqm_s^X(A) + p(1 - q)m_s^X(\bar{A}) \text{ , для } A \neq X \\ m^X(X) = pqm_s^X(X) + p(1 - q)m_s^X(\emptyset) + 1 - p \text{ , для } A = X. \end{array} \right. \quad (6)$$



ЕКСПЕРИМЕНТАЛЬНЕ ДОСЛІДЖЕННЯ



Об'єкт інтересу може перебувати в одному з 4 станів $\{a, b, c, d\}$.

Два джерела даних: 1) розподіл гіпотез $\{0,3;0,15;0,25;0,3\}$;

релевантність – $p=0,7$; достовірність – $q=0,8$.

2) розподіл гіпотез $\{0,2;0,3;0,3;0,2\}$;

релевантність – $p=0,5$; достовірність – $q=0,7$.

Правила комбінування

	Демпстера	Дисконтування	Сметса	Жанга	З врахуван. метаданих
Результат ранжування гіпотез	$a \rightarrow c \rightarrow b \rightarrow d$	$a \rightarrow c \rightarrow b \rightarrow d$	$a \rightarrow c \rightarrow b \rightarrow d$	$a \rightarrow c \rightarrow b \rightarrow d$	<u>$c \rightarrow a \rightarrow b \rightarrow d$</u>



ВИСНОВКИ



- Висвітлені проблеми вирішення тематичних задач ДЗЗ з використанням різних джерел, серед яких: великі обсяги даних ДЗЗ, обмежений ресурс часу на їх оброблення та доставку користувачам, тенденції щодо уніфікації форм подання інформації, наявні різнотипні дані джерел інформації, зростаючі вимоги до достовірності та повноти інформації ДЗЗ і обмежені можливості кожного з відомих засобів.
- Розглянуті основні та альтернативні правила комбінування джерел в теорії свідчень, запропоновано врахувати метадані про джерела (релевантність і достовірність) в розрахунках.
- Запропонований підхід до моделювання процесів вирішення тематичних задач ДЗЗ на основі теорії свідчень дозволяє враховувати характеристики джерел при їх комбінуванні.



КОНТАКТИ



ДЯКУЄМО ЗА УВАГУ!

- Попов Михайло Олексійович
- Науковий Центр аерокосмічних досліджень Землі НАН України
- Україна, 01601 Київ,
- ул. Олеся Гончара, 55-Б
- Тел/факс (38044) 482-01-66 mpopov@casre.kiev.ua

- Зайцев Олександр Вікторович
- Воєнно-дипломатична академія
- Україна, 01601 Київ,
- ул. Мельникова, 81
- Тел/факс (38044) 483-15-05 a.zaysev@gmail.com