

**Национална Научна Програма Информационни и
Комуникационни Технологии в Науката, Образованието и
Сигурността**

НАУЧЕН ОТЧЕТ

**Компонент 1: Електронна инфраструктура за отворена наука
и отворен достъп до научни резултати**

Основни цели:

1. Развитие на Центровете за високопроизводителни и разпределени пресмятания на участващите институции
2. Създаване на математически модели, високоефективни алгоритми и софтуер за решаване на важни научни, научно-приложни и технологични задачи от различни области
3. Създаване и поддръжка на до 15 пилотни хранилища, осигуряващи свободен достъп до научни публикации и научни резултати чрез системи за лицензиране.

Дейностите за постигане на горните цели са разпределени в четири работни пакета:

- 1.1. Високопроизводителни и разпределени пресмятания (р-л доц. д-р Емануил Атанасов от ИИКТ-БАН) с една научна задача.
- 1.2. Компютърно и математическо моделиране с приложение в инженерните и природните науки (р-л проф. дмн Петър Бойваленков от ИМИ-БАН) с осем научни задачи.
- 1.3. Хранилища за отворени научни резултати (р-л проф. дмн Петър Станчев от ИМИ-БАН) с една научна задача.
- 1.4. Съхранение и анализ на големи обеми данни и технологии за приложение на изкуствен интелект (р-л проф. дфзн Леандър Литов от СУ) с една научна задача.

Работен пакет 1.1., Научна задача 1.1.1. Интегриране на съвременните изчислителни системи и системи за съхранение на данни, софтуер, мидълуер и услуги; Дейности по включване на центровете за високопроизводителни изчисления в европейски и национални облачни системи за отворена наука. Извършени са дейности по интегриране на наличните съвременни изчислителни системи с цел ефективно им предоставяне на български изследователи. Инсталирани са и са тествани нови научни приложения и средства за разработка, проведени са обучения на млади учени. Създадена е методология и са проведени тестове на контейнерно-базиран метод за стартиране на MPI приложения, които да подпомогнат оптимизацията на приложенията и насочването на потребителските групи към най-подходящите за тях изчислителни ресурси. По този начин ще се подпомогне синергията между групи със сходни задачи, което ще бъде акцентът през следващите години на програмата.

Получените резултати са докладвани на научни конференции и са включени в 6 публикации с SJR.

Работен пакет 1.2 включва 8 задачи, свързани с различни области на приложение в науката и практиката на математически методи и модели.

Научна задача 1.2.1. Разработване на *in silico* методи и инструменти за изследване на комплексни среди в контекста на иновативни технологични решения. Представено е proof-of-concept изследване за разработване на иновативен метод за 3D визуализация, базиран на съществуване на две независими вътрешни геометрични структури в протеините – породени съответно от пептидните равнини и страничните вериги. Получените експериментални резултати са подкрепени от биологичните данни и предоставят нова перспектива за бъдещото разбиране на структурата и функцията на хистадина в миоглобина. Разработени са нови ефективни числени алгоритми (адаптирани Primal-Dual) за обезшумяване, възстановяване и сегментиране на изображения. Направени са теоретични и експериментални сравнения на дробен оператор на Лаплас върху крайна област с подходящи гранични условия. Анализирани са числената ефективност на йерархичен алгоритъм за решаване на системи линейни алгебрични уравнения, възникващи при крайно-елементни дискретизации на дробни дифузионни процеси. Изследвани са рискове при работа с компрометиранни данни. Изследвани са връзките между набор от клинични параметри с прилагане на йерархичен клъстерен анализ.

Изследвана е селекция от 1400 хранителни протеини чрез PLS-DA след преобразуване на основните параметри в еднообразни вектори подбрани на базата на входната матрица. Приложен е Клъстерен анализ, за да се постигне определено ниво на разделяне в малък набор от растителни протеини. Предложен е лесен метод за модифициране чрез топковото смилане приложено за определяне на фотокаталитичните свойства на хумини, получени от H₂SO₄ и дехидратиран глюкозо-фруктозен сироп. Предложена е изчислителна стратегия за предсказване на разпределението за набор от разтворители при съответно фазово равновесие. Показана е схема за моделиране на качеството на почвата чрез помощта на интелигентен анализ на данни. С помощта на статистически модели за сложни система, като например моделиране на попадането на наночастици в околната среда, с цел определяне на връзките между физикохимичните параметри на наночастиците със съответните стойности на токсичност.

Получените резултати са докладвани на научни конференции и са включени в 9 публикувани или приети за печат публикации с импакт фактор и/или SJR.

Научна задача 1.2.2. Моделиране и оптимизация на моно- и би-метални нано-материали с желани свойства както и нови по-ефективни материали за наноелектрониката, системите за съхранение и преобразуване на енергия. Направен е обзор на съществуващи подходи за моделиране на микроструктури. Създаден е хибриден стохастичен метод (тип Монте Карло) за откриване на устойчиви клъстерни конфигурации от наночастици. Методът е приложен и изследван за устойчивост при изследване на структури от малък брой атоми. Получените резултати са докладвани на научни конференции и са включени в 2 публикации с импакт фактор и/или SJR.

Научна задача 1.2.3. Разработване на нови математически методи и изчислителни техники за изследване и валидиране на динамични модели, описани чрез обикновени и/или частни диференциални уравнения с приложение. Разработена е и е приложена нова компактна диференчна схема с четвърти ред на точност по пространствената променлива за системи от параболични частни диференциални уравнения със свързани нелинейни реакции. Разработен е нов алгоритъм на Монте Карло за решаване на системи от линейни алгебрични уравнения, основан на наскоро разработения метод Монте Карло „случайно блуждаене по уравненията“. Разгледани са многомерни интегрални от два типа, намиращи приложение в Бейсовската статистика, чрез подобрен квази-Монте Карло метод, базиран на квази-случайната редица на Собол с побитово отместване и разбъркване. Направено е приближено пресмятане на пълните индекси на чувствителността посредством квази-случайната редица на Собол и Фор. Стохастични алгоритми базирани на подобрената редица на Собол са

приложени за пресмятане на индекси на чувствителността, съответстващи на шест химични реакции и четири групи различни замърсителя.

Изследвани са в няколко сценария реакциите на моделиращите метаматериали Miura-ori за динамично компресиране в равнината. Предложен е математически модел на имунния отговор при заразяване с вируса денга. Предложен е общ подход за стабилизиране на динамиката на двумерен нелинеен модел на хемостат (биореактор), описващ анаеробно разграждане на органични отпадъци в замърсени индустриални и градски води. Предложен е математически модел на биореактор за биологично разграждане в отпадни води, съдържащи смес от химически замърсители фенол и р-крезол.

Продължено е изследването на иновативен алгебричен интервален модел на линейни уравнения на равновесие, които са в основата на крайно-елементни модели за анализ и дизайн на строителни конструкции с интервални моделни параметри. Разработен е и е изследван нов метод за получаване на параметризирано решение на линейна алгебрична система, съдържаща линейни ранк-едно зависимости между интервално-значни параметри.

Изследвани са решенията на линейни кооперативни, слабо-свързани системи елиптични уравнения. С помощта на принципа за субординация се изучава фундаменталното решение на многомерното уравнение на дифузия, съдържащо дробни производни по времето и пространството. Разглежда се задачата на Коши за еволюционно уравнение, съдържащо оператори от дробен ред по времето и по пространството. Разгледана е задача за нано-пукнатина в магнито-електро-еластична равнина (МЕЕ) подложена на натоварване, получено от срязваща вълна, като симулациите показват зависимостта на концентрация на напрежението на динамичното поле от свързаността на материала, честотата на падащата вълна, размера на нано-пукнатината и повърхностните материални свойства. Разгледано е първото собствено число ламбда на р-Лапласиана с условия на Дирихле в ограничена област на n-мерното пространство, като е получена аналитична оценка отдолу на ламбда чрез неравенство на Харди с двойно сингулярно ядро. Изследвано е глобалното поведение на слабите решения на двойно-дисперсни уравнения от шести ред с линейна възстановяваща сила, като разглежданите задачи моделират физически процеси – разпространение на напречни вълни в изотропен цилиндричен прът, разпространение на вълни с ненулево тензорно напрежение и др. Изучава се осцилиращ Ненютонов флуид в безкраен канал, като задачата се появява при изследване на различни физиологични, химични и биохимични процеси, като например кръвен или полимерен флуид. Направен е критичен анализ и обзор на съществуващите в литературата фундаментални решения на уравненията на еластодинамиката при решаване на тримерни динамични задачи, както и за функции на Грийн за класове от тримерни динамични задачи за различни материали. Предложен е модел за задача за вълново разсейване и дифракция на срязващи вълни от нано-пукнатина в магнито-електро-еластична равнина (МЕЕ), който отчита повърхностните свойства, размера на нано-пукнатината и взаимодействието между пукнатината, падащата вълна и много-функционалната МЕЕ матрица.

Анализирано е неотдавнашно разширение на т.нар. модифициран метод на най-простото уравнение. Обсъжда се последната версия на метод за получаване на точни решения на нелинейни частни диференциални уравнения, като се показва, че редица досегашни методи са частни случаи. Изследвани са груповите свойства на системата от уравнения, която описва обобщен случай на двустенна въглеродна нанотръба, а именно съответните коефициенти трябва да зависят от пространствената променлива. Реакциите на моделиращите метаматериали Miura-ori за динамично компресиране в равнината са изследвани за няколко сценария, а именно: компресиране на метаматериали с еднаква плътност при постоянна скорост и импулсно натоварване, моделирано като въздействие с начална скорост и масово въздействие на метаматериал със степенувана плътност.

Получените резултати са докладвани на научни конференции и са включени в 8 излезли, 12 приети и 3 предложени за публикуване статии. Всички списания имат SJR и/или импакт

фактор. Публикуваната статия е с импакт фактор и е отчетена частично и в Задача 1.2.2.

Научна задача 1.2.4. Модели за анализ на данни и откриване на закономерности от данни при оценка и предсказване на безопасността или токсичността на лекарства.

Създаден е масив от данни, който да играе ролята на основен информационен източник за целите на изграждането на планираната база от данни за човешки туморни антигени. Разработен е подход за търсене на доказани човешки туморни антигени, свързани с различни заболявания – по ключови думи, за анотирани и рецензирани протеини, пряко свързани с туморни антигени при човека. Разработен е и е тестван модел за предсказване на имуногенност на ТАг въз основа на първичната структура, основан на метода на хомоложно (еволуционно) моделиране на протеини. Изследването е фокусирано и върху изчислително моделиране на вирусоподобни частици на хепатит В с вложки от епитопа на хепатит Е, като е предложена стратегия за избор на оптимално място за вмъкване на допълнителната последователност чрез моделиране на вариации в позицията на вложката и сравняването им с нативния протеин, използвайки структурно подравняване.

Получените досега резултати са включени в статия, която е предложена за публикуване в списание с импакт фактор.

Научна задача 1.2.5. Предвиждане и гарантиране на качеството в човеко-кибер-физически системи.

Предложени са мащабируеми концептуални модели за представяне на човеко-кибер-физически системи, посредством виртуални обслужващи устройства. Изведени са уравнения за описване на натоварването на такива системи, в които има опашки, където заявките чакат за обслужване. В изследваните цялостни обслужващи системи, продължителността на обслужване в сървъра зависи и от продължителността на чакане в опашката, което е слабо изследвана зависимост. Анализирани са средствата за концептуално моделиране на цялостни обслужващи системи, което включва и новия подход на Обобщените мрежи. Направени са предложения за класификация на средствата и развитие на методологията за концептуално моделиране на обслужващи системи. Разработват се нови числени методи за изследване на чувствителността (посредством изучаване на производните) на модели на сложни системи, каквито са човеко-кибер-физическите.

Проучени са някои епидемиологични модели, които в някои ситуации могат да бъдат приложени към теорията на разпространението на компютърните вируси. Изследват се кооперативното разпределение на политиките за филтриране на трафика и процентът на трафика, обяснен с автоматично генерирани идентификатори. Изследвани са софтуерни инструменти за разпределени бизнес приложения. Изучени са някои нови свойства на съществуващи параметрични фамилии от сигмоидални функции използвани в областта на “размити множества“. Получени са и интересни апроксимационни резултати с приложение в невронни мрежи.

Получените досега резултати са докладвани на научни конференции и са включени в 21 публикации с импакт фактор и/или SJR. В процес на изграждане е международна научна мрежа с участие на ИМИ-БАН, ЮЗУ Благоевград; Университета на Виена, Австрия; Национален изследователски университет „Висша школа по икономика“, Перм, Русия, Лаборатория по информационни технологии на Обединения институт за ядрени изследвания в Дубна, Русия, и други.

Научна задача 1.2.6. Конструирание, класификация и анализ на комбинаторни структури.

Предложен е метод за конструирание на шумозащитни линейни кодове базирани на wavelet трансформация и на AMD (Algebraic Manipulation Detection) кодове, получени от тях, с възможни приложения в преносими устройства за памет. Получени са нови класификационни резултати (границы) за мощността на кодове с фиксирано минимално и максимално разстояние между кодовите думи и за потенциалната енергия на кодове със зададено максимално разстояние и фиксирана мощност. Изследвани са кодове с две разстояния (последователни числа), които апроксимират кодове с ограничена енергия за амплитудно-фазова модулация. Анализирани са DLT (Distributed Ledger Technology) с

различни характеристики и е изследвана тяхната пригодност за решаване на текущи проблеми в областта на управлението на жизнения цикъл на софтуера. Показано е, че прилагането на различни пермутации в различните рундове на обобщена схема на Файстел може да доведе до подобряване на дифузията на съответния шифър. Предложен е обобщен свиващ-мултиплексиращ генератор, управляван от троична m -редица GSMG-3m (Generalized Shrinking Multiplexing Generator controlled by ternary m -sequences).

Получените досега резултати са приети за докладване на научни конференции и са включени в 7 публикации с SJR.

Научна задача 1.2.7. Моделиране, дизайн и управление на биомехатронни устройства и системи. Разработен е хибриден подход на паралелно задвижване в ставите на активен екзоскелетон на горен крайник, включващ пневматични актуатори и ниско инерционен електромотор. Проведени са компютърни симулации за оценка на свойствата на екзоскелетона да осигури силова рефлексия. Получени са оценки на вътреставни движения в колянната става по време на активно изометрично разтягане на мускулните групи на долния десен крайник. Започната е работа по проектиране и изработване на мехатронни системи за рехабилитация, които да са насочени към индивидуалния потребител, или групи от потребители с подобни физиологични характеристики.

Получените досега резултати са публикувани в 3 работи, 1 от които с SJR.

Научна задача 1.2.8. Софтуерни инструменти, системи и модели за проектиране на мултиплатформена Операторска станция. Направен е анализ на съществуващи реализации на Операторски станции. Подготвена е функционална спецификация на универсална Операторска станция и се работи върху концептуален модел на контролиран обект, както по проектиране на хетерогенни жично-безжични мрежи. Направена е заявка за два Полезни Модела в Патентното Ведомство - Устройство за управление на прожекцията на лазерни анимации и Електронен контролер за предаване на сензорни данни. Към реализацията на задачата са привлечени 4 бизнес-партньора.

Получените досега резултати са приети за докладване на научни конференции и са включени в 2 излезли публикации с SJR.

Работен пакет 1.3. Хранилища за отворени научни резултати, Научна задача 1.3.1.

Създаване на модел, софтуерни платформи и цифрови хранилища за съхраняване и публикуване на научни резултати с цел въвеждане в България на идеите, практиките и политиките на Европейския съюз свързани с отворената наука и присъединяване на България към EOSC (European Open Science Cloud). В две аналитични изследвания са проучени най-добри практики, отворени стандарти и свободен софтуер за реализация на цифрови хранилища, съхраняващи научни резултати с отворен достъп в EOSC. Направен е подробен анализ на различни информационни хранилища за съхраняване на научни резултати. Разработена е и е предложена цялостна визия и множество от дейности и конкретни стъпки за присъединяване на България към EOSC. Проектирани са софтуерна и хардуерна архитектура за национален и институционален репозиториум за съхраняване на резултати от научна дейност. Създаден е модел на метаданните за описание на всички резултати от научната дейност съвместим с OpenAIRE и CRIS. Извършена е подготовка за създаване на цифрови хранилища за съхраняване и публикуване на научните резултати на ТУ-Варна, ХТМУ-София и УНИБИТ.

Получените досега резултати са докладвани на научни конференции и са включени в 4 излезли публикации с SJR в издания с отворен достъп.

Работен пакет 1.4. Съхранение и анализ на големи обеми данни и технологии за приложение на изкуствен интелект. Включени са дейности, свързани с информационното обслужване на научни изследвания, генериращи големи обеми от данни (от порядъка на PB) включващо съхранение, анализ и оценка на тяхното качество. Работи се по следните

подзадачи: развитие на центъра за Grid и облачни технологии на СУ „Св. Климент Охридски“, изграждане на преден интерфейс към голяма международна облачна инфраструктура, инсталиране на нови мощности и интегрирането им в световната Grid мрежа, разработване на методи и технологии базирани на изкуствен интелект за off-line контрол и оценка на качеството на работата на апаратура за данни от експеримента CMS. Тестван е ML софтуера, разработен за DT и RPC подсистемите на експеримента CMS, като естествената следваща стъпка е разширяването на софтуера до пълната Мюонна система на CMS, а евентуално в последствие и до други системи на експеримента. Използван е клъстерът LXPLUS в ЦЕРН, NPC инфраструктурата на IDRIS и по-конкретно суперкомпютъра Ouessant <http://www.idris.fr/ouessant/> и облачната инфраструктура на ЦЕРН. Преработеният софтуер е публикуван и достъпен от следното хранилище:

https://github.com/bapavlov/L1MTriggerRateMonitoringWithML/tree/rpc_v1/python

Работата по пакета се извършва в тясно сътрудничество в ЦЕРН и колаборацията CMS.

Задачите, техните цели и постигнати през първите 6 месеца резултати са описани подробно в приложените отчети по задачи, където са посочени и количествените показатели за изпълнение на програмата (по задачи). В отчитаните публикации е отбелязан по различни начини приносът на Програмата.

Обобщена информация за основните получени резултати и изпълнени индикатори (следвайки структурата на Анекс 1 към тези указания) е представена в следната таблица.

Показатели	Общо
Брой научни публикации в специализирани списания и/или научни поредици с импакт-фактор (IF) и/или импакт-ранг (SJR)	73
Брой разработени и предложени тримерни модели	0
Брой участия в национални и международни научни форуми и изложения	59
Брой мероприятия за популяризиране на получените резултати (конференции, семинари, кръгли маси, информационни дни)	13
Брой на млади учени, докторанти и студенти, участвали в програмата	32
Брой изградени международни научни мрежи	4
Брой бизнес партньори, привлечени в изпълнение на програмата	4
Брой браншови и други организации, привлечени за изпълнение на програмата	1
Брой заявени/регистрирани полезни модели	2

10. ПОКАЗАТЕЛИ ЗА ИЗПЪЛНЕНИЕ НА ПРОГРАМАТА И ИНДИКАТИВНИ СТОЙНОСТИ

а) брой потребителски групи за ползване на е-инфраструктура	9
б) брой подпомогнати успешни интердисциплинарни пилотни (до 3-6 месеца) проекти	1
в) брой изградени и/или разширени хранилища за отворена наука	4(в процес на изграждане)
г) брой създадени научни мрежи	3
д) брой подадени съвместни научни проекти	3
е) брой високотехнологични публикации в областта на програмата в списания в ТОП 10% в класацията на Thompson-Reuters WoS	0
ж) брой обучени млади изследователи	20

Изготвил отчета:

Проф. дмн Петър Бойваленков (ИМИ-БАН)
Ръководител Компонент 1

