

схему дій стосовно слідчої ситуації. Він має обмежену функцію, оскільки працює тільки у визначеній ситуації. Разом з тим його позитивна роль полягає у швидкості вирішення поставленого завдання, що певною мірою сприяє ефективності розслідування. У нетрадиційних, нетипових ситуаціях, що вимагають творчого пошуку і підходу, алгоритм може грати допоміжну, орієнтуючу роль. Розроблені алгоритми слідчих дій, як і їхні системи, повинні складати банки даних, що мають за мету інформаційне забезпечення розслідування. В цьому плані оптимальні системи слідчих дій можуть виконувати двояку роль: а) навчальну; б) практичного використання. Навчальна роль полягає в ознайомленні із системами слідчих дій, спрямованих на ефективні вирішення завдань розслідування. Практичне значення полягає в реалізації названих систем у процесі розслідування як забезпечуючих ефективність і швидкість стосовно ситуацій злочинів, складних для розслідування, — легалізація грошових коштів, одержаних злочинним шляхом, терористичні акти, бандитизм, злочини, вчинювані організованими групами, що сприятиме ефективності розслідування, виконуючи навчальні і реалізовані в розслідуванні функції.

Надійшла до редколегії 10. 10. 05

О. Юрченко, кандидат юридичних наук
(Національна академія Служби безпеки
України)

Можливості застосування новітніх технологій в інформаційно-аналітичному забезпеченні державної системи захисту промислової власності

З проголошенням Урядом України курсу на створення ринкової економіки, вибором інноваційної моделі розвитку, правовим закріпленням різноманітних форм і видів власності, визнанням об'єктів інтелектуальної, в тому числі промислової, власності як товарів зростає нагальна потреба в їх ефективному використанні та захисті від злочинних посягань. У нових умовах назріла об'єктивна необхідність у принциповому поновленні й вдосконаленні системи захисту промислової власності¹. Значну роль у вирішенні цих

¹ Див.: Юрченко А. М. Проблемы защиты интеллектуальной собственности Украины в контексте евроинтеграционного развития // Бизнес и безопасность. – 2004. – № 4 (41). – С. 5–7.

питань держава відводить Службі безпеки України, зокрема підрозділам контррозвідального захисту інтелектуальної власності.

Вирішення цих завдань на сучасному етапі розвитку науково-технічного прогресу неможливо без використання новітніх інформаційних технологій. Тому розроблення та впровадження інтелектуальних систем аналізу й оброблення інформації в контррозвідальну діяльність СБУ є актуальними чинниками підвищення ефективності оперативно-службової діяльності.

Питанням інформаційно-аналітичного забезпечення оперативно-службової діяльності правоохоронних органів присвячено праці вітчизняних науковців: С. Азарова, М. Вергузаєва, В. Голубєва, Р. Калюжного, М. Костюченка, В. Цимбалюка, В. Хахановського, В. Хорошка, М. Швеця та ін. Проте публікації щодо впровадження інтелектуальних інформаційних систем нового покоління в діяльність правоохоронних органів України практично відсутні.

Аналіз сучасних технологій оброблення даних в інтересах державних, правоохоронних і комерційних структур свідчить про активні розроблення і використання технологій інтелектуального аналізу даних (ТІАД) і геоінформаційних систем (ГІС-технологій). Багато в чому це пов'язано з тим, що на застосуванні зазначених технологій оброблення даних базуються перспективні концепції управління силами і засобами в складній оперативній обстановці. На основі їх активного використання приймаються складні політичні, економічні й інші рішення.

З розвитком Internet-технологій багато провідних зарубіжних фірм, що спеціалізуються в сфері інформаційних технологій, щорічно вкладають величезні кошти в створення відповідного інструментарію інтелектуальної інформації. Причому навіть самі фантастичні вкладення тут виправдані — вже через три роки, на думку глави Intel К. Буррета, на Землі прогнозується мільярд користувачів Internet.

Світовий Internet став настільки глобальним, що боротьба зі злочинами в ньому вже не під силу навіть такому монстру, як ФБР¹. Американська спецслужба розробила спеціальну комп'ютерну програму InfraGuard для залучення різних організацій і бізнес-структур, у тому числі іноземних, у справу виявлення і притягнення до відповідальності кіберзлочинців. На сьогодні програму встановлено в офісах близько 500 компаній, включаючи такого гіганта, як IBM. Безумовно, інтелектуалізація оброблення даних у зв'язку з бурхливим розвитком Internet-технологій починає здобувати все більше значення для ефективного розв'язання двох основних проблем.

¹ Див.: Безопаска комп'ютерних систем. Комп'ютерна злочинність та її попередження / М. С. Вергузаєв, В. О. Голубєв, О. І. Котляревський, О. М. Юрченко / Під ред. О. П. Снігерьова. – Запоріжжя, 1998.

З одного боку, це одержання в реальному масштабі часу багатоаспектної й об'єктивної інформації з Internet, інших доступних інформаційних і телекомунікаційних систем про стан, напрямок розвитку і рівень погроз тих чи інших процесів у світовому масштабі, суспільстві та його окремих етно-економічних утвореннях. Тут зовсім очевидно, що синтез знань, відображених у тих чи інших розрізненних і, на перший погляд, не пов'язаних єдиною логікою даних, здійснений на основі систем їх інтелектуального оброблення, дає абсолютно нову інтегративну якість, що дозволяє вгадати, а відтак, і попередити негативний розвиток тих чи інших процесів і явищ. Таким чином, йдеться про створення на базі названих технологій інтелектуального середовища моделювання безпечного розвитку систем і процесів¹.

З іншого боку, застосування інтелектуальних технологій оброблення даних дає можливість на порядки підвищити ефективність функціонування різних комп'ютеризованих систем, у тому числі пов'язаних із прийняттям стратегічних рішень, що відповідають за безпечний розвиток технологічних процесів, використовуються у системах управління військами і зброєю, правоохоронними структурами.

З наведеного можна зробити висновок: в умовах обстановки, що швидко змінюється в Україні і характеризується вираженим дефіцитом часу на прийняття рішень, пов'язаних із глибоким аналізом і прогнозуванням розвитку ситуації, орієнтація на інтелектуальні автоматизовані системи підготовки рішень виступає першочерговим завданням.

В умовах стрімкого збільшення кількості електронних записів ТІАД і ГІС-технології потрібні більшості користувачів ними. Ці технології, створені зусиллями провідних світових виробників, стали визнаним стандартом для низки міжнародних організацій. У чому ж полягають їх сутність, привабливість і корисність?

Якщо коротко, сутність цих технологій полягає в інтелектуалізації управління процесами інформаційного забезпечення, що включає автоматизований пошук інформаційних взаємозв'язків будь-яких об'єктів: подій, людей, формалізованих знань із усіх видів баз даних, що існують у світі, у тому числі баз даних неструктурованої текстової інформації, візуалізацію складно структурованої інформації в графічному і табличному вигляді, статистичний і географічний аналіз подій, пошук прихованих, латентних об'єктів.

¹ Див.: Системно-структурный подход к совершенствованию аналитической работы в органах внутренних дел в условиях создания, функционирования и развития ОАСУ (ЕИВС) / Ю. И. Аболонцев, М. С. Вертузаев, В. А. Гадышев, Т. А. Мацко: Учеб. пособ. – М., 1992.

Говорячи про ТІАД, слід зазначити, що вони розраховані на користувача з мінімальною комп'ютерною підготовкою. Робота в інтелектуальному середовищі здійснюється в реальному масштабі часу. Будучи винятково простими і надійними в застосуванні, технології забезпечують істотну економію витрат і часу на пошук значущої інформації. Особливо ефективно вони зарекомендували себе за рубежом при практичному плануванні заходів у сфері боротьби з організованою злочинністю, незаконним обігом наркотиків, економічними злочинами¹.

З огляду на виняткову важливість антикримінальної діяльності розглянемо основні напрямки ефективного застосування ТІАД у цій сфері.

По-перше, це розслідування фактів шахрайства². ТІАД дають можливість оперативно збирати, глибоко аналізувати інформацію з різних джерел (повідомлення, результати попереднього розслідування, інформація про зв'язки сторін, адреса, банківська і фінансова інформація), встановлювати неявні зв'язки. Крім того, технології дозволяють проводити часовий аналіз, виявляти нестиковки в подіях, виводячи експерта на факти шахрайства з указівкою місцезнаходження грошей і рахунків.

По-друге, ТІАД різко підвищують продуктивність роботи слідчих при розслідуванні масштабних справ, діяльності організованих злочинних груп із десятками і сотнями фігурантів, тисячами епізодів і зв'язків, злочинної активності на значних територіях; систематизують події, впорядковують їх за типами і значущістю, структурують за часом, визначають мету, мотивацію і структурну організацію злочинців у тому чи іншому кримінальному епізоді.

По-третє, ТІАД є незамінними в так званому розвідувальному аналізі³, а саме: на етапі планування при визначенні мети, робочих гіпотез, корисній інформації й організації збирання даних та їх тестуванні, потім — на етапі інтерпретації даних і поданні результатів у вигляді діаграм, схем, таблиць, графіків.

По-четверте, ТІАД показали свою сильну профілактичну сторону в процесі визначення потенційних об'єктів і суб'єктів кримінальної ак-

¹ Див.: Вертузаев М. С., Попов А. Ф., Юрченко А. М. Мошенничество с использованием пластиковых платежных средств (из опыта работы ГУВД г. Москвы) // Бюллетень по обмену опытом работы органов внутренних дел Украины. – К., 1998. – № 122. – С. 49–54.

² Див.: Шахрайство з пластиковими платіжними засобами. / М. С. Вертузаев, Я. Ю. Кондратьев, С. Є. Пугачов, О. М. Юрченко // Бюллетень з обміну досвідом роботи. – К., 2000. – № 125. – С. 14–17 (вих. № 28/12т); № 126. – С. 22–29 (вих. № 28/28цт, інв. № 9–369 НАВСУ); № 127. – С. 23–26 (вих. № 28/42цт, інв. № 9-372 НАВСУ).

³ Див.: Вертузаев М. С., Вертузаев А. М., Юрченко А. М. Некоторые аспекты виртуальной разведки // Бизнес и безопасность. – 2002. – № 6. – С. 54–58.

тивності, наприклад ідентифікація потенційних злочинців¹, з'ясування цілей, часу й об'єктів можливого злочину, запобігання масовим злочинам і терористичним актам, прогнозування можливостей і напрямків промислового шпигунства, виявлення взаємодії злочинних елементів у всьому ланцюжку незаконного обігу наркотиків.

Нарешті, не можна не згадати і про таку можливість технологій, як створення оптимальних схем організації роботи оперативно-слідчих груп з огляду на особливості «кримінальної моделі», яка відпрацьовується.

Щоб розмова не була лише абстрактною, розглянемо більш докладно один з найбільш відомих програмних продуктів класу — i2, який звучить як «Ай-Ту». Технології «Ай-Ту»:

- використовують понад 1300 державних і комерційних організацій у 70 країнах світу;
- прийняті як стандарт Інтерполу й Європолу;
- відзначені спеціальними нагородами Міжнародної асоціації аналітиків правоохоронних органів (IALEIA — International Association of Law Enforcement Intelligence Analysts) у 1993, 1998, 2001 pp.;
- прості в застосуванні, не вимагають від користувача спеціальних комп'ютерних знань.

Analyst's Notebook (ANB) — основний компонент технологій «Ай-Ту», забезпечує переклад і демонстрацію аналітикам складної інформації у вигляді схеми, що легко розуміється. При цьому інструментальний засіб Link Notebook підтримує схеми аналізу зв'язків, потоків продукції, фінансових потоків і т. ін., у той час як Case Notebook «відповідає» за часові графіки чи схеми послідовності подій, діаграми дій у кожній з подій, комплексні діаграми, що відображають події та їх перебіг.

Таким чином, ANB може автоматично проаналізувати всі знайдені дані, розплутати складні зв'язки у великих мережах, що відображають взаємодії об'єктів різної природи, і швидко представити все це в інтелектуальній графічній формі, полегшивши роботу аналітика і різко просунувши його до вирішення поставлених завдань, наприклад, при розкритті і розслідуванні злочинів.

iBase — система управління даними — становить найбільший інтерес для користувачів, зацікавлених у створенні власної бази даних. Вона дозволяє збирати, структурувати, зберігати й оновлювати дані з різнорідних джерел, складає звіти й обробляє дані за допомогою системи візуальних запитів за принципом «намалюй питання — одержи картинку-відповідь».

¹ Див.: Вертузаев М. С. Специфические проблемы идентификации личности при обработке информации на ЭВМ // Технические средства и системы в предупреждении и раскрытии преступлений. — К., 1991. — С. 78–84.

Система управління даними iBase включає спеціальні інструментальні засоби накопичення й оброблення складно структурованої інформації з розширеними можливостями по взаємодії з графічними засобами аналізу, відрізняючись, проте, інтуїтивно зрозумілим інтерфейсом, що виключає необхідність вивчення складної мови запитів бази даних¹.

Важливим для практики є те, що iBase надає можливість працювати «у команді», тобто в реальному масштабі часу поділитися з колегами знаннями про ту чи іншу ситуацію, що виникає, наприклад, у процесі розкриття і розслідування злочинів. При цьому, зрозуміло, забезпечується повна інтеграція з ANB.

iBridge — інтерфейсні програмні засоби зв'язку існуючих баз даних і додатків з Analysts Notebook:

- у сполученні з Analyst's Notebook становлять могутнє й ефективне аналітичне середовище;
 - дозволяють швидко витягти й об'єднати в одній діаграмі інформацію з усіх доступних баз даних;
 - ефективно виявляють приховані взаємозв'язки в існуючих даних.
- Analyst's Workstation:
- поєднує можливості Analyst's Notebook, iBase, iGlass (компонента, що полегшує побудову графіків);

– дозволяє інтегруватися з ПС-технологіями й іншими технологіями аналітичного оброблення даних. До речі, пакет iGlass надає різноманітні можливості для аналізу даних, використовуючи комбінацію запитів, графіків і діаграм. За його допомогою оперативно виділяються і порівнюються потрібні схеми і взаємозв'язки між ними. При цьому унікальна навігаційна модель у iGlass дає можливість перевіряти різні гіпотези, створювати діаграми в різноманітних форматах і стилях, просуюючи дослідницьку ідею до оптимального втілення.

Виробник технологій «Ай-Ту» компанія i2.Ltd заснована в 1990 р., є лідером світового ринку програмних інструментів для візуалізації й аналітичного оброблення інформації.

Ексклюзивним дистриб'ютором технологій «Ай-Ту» у країнах СНД виступає компанія РДТЕХ (заснована в 1992 р.). Це відомий системний інтегратор у сфері проектування, розроблення, впровадження і супроводу інформаційних систем. Нею здійснено низку великих розробок в інтересах правоохоронних структур.

Технології «Ай-Ту» мають надзвичайно широкі сфери застосування. Їх використовують для інформаційно-аналітичного забезпечення держав-

¹ Див.: Вертузаев М. С., Цыганов О. Г. Построение интеллектуального интерфейса сотрудников следственных и оперативных аппаратов с АБД // Компьютерные технологии в правоохранительной сфере: — М., 1993. — С. 84–93.

них структур при вирішенні завдань вдосконалення системи національної безпеки, служби охорони вищих посадових осіб країни, військової розвідки, безпеки закордонних посольств і представництв, митної діяльності, підрозділів по податках і зборам, податкової поліції, контролю в сфері імміграційної служби, органів виконання покарань, судочинства.

Технології приносять велику користь у діяльності аналітичних служб правоохоронних структур при організації ефективних заходів щодо боротьби зі злочинністю, а саме: при розслідуванні складних багатоепізодних злочинів, розкритті серійних злочинів, боротьбі з організованою злочинністю, незаконним обігом наркотиків і зброї, шахрайством, виробництвом і поширенням контрафактної продукції, «відмиванням» грошей, злочинами в сфері високих технологій, автомобільного бізнесу, боротьбі з тероризмом, при організації безпеки масових заходів, пошуку зниклих і т. под.

«Ай-Ту» має цікаві бізнес-додатки. На основі названих технологій вирішуються завдання забезпечення ділової розвідки, безпеки банківської діяльності, корпоративної безпеки, діяльності з попередження шахрайства в сфері кредитних карт, страхового бізнесу, телефонних мережах і телекомунікаційних системах. Деякі відомі журналісти застосовують «Ай-Ту» для проведення незалежних розслідувань.

Отже, наявні практичні приклади застосування технологій «Ай-Ту» наочно підтверджують ту думку, згідно з якою перспективи сучасних аналітичних технологій найчастіше обмежені лише фантазією дослідника.

Останнім часом технології «Ай-Ту» стали активно застосовуватися за рубежом у сполученні з ГІС-технологіями. ГІС (геоінформаційні системи) — це клас програмних продуктів, основним завданням яких є оброблення й аналіз даних, що мають просторову локалізацію. До таких даних належить будь-як інформація, що містить, крім описової (атрибутивної) складової, координатну компоненту.

Поняття ГІС-технології набагато ширше за «програмне забезпечення для геології, географії, картографії». Застосування технологій ГІС надає різноманітні можливості з оброблення як цифрових карт чи планів місцевості, так і планів будинків, комунікацій, дозволяє формувати складні комбіновані запити до даних, що включають низку просторових обмежень.

Можна виділити такі переваги використання ГІС при інтелектуальному аналізі даних:

1) візуалізація складно структурованої інформації. Цифрова чи карта-план дозволяє експерту оцінити спільне розташування цікавлячих його об'єктів та їх топологічний взаємозв'язок;

2) можливість вибірки інформації з урахуванням просторових обмежень, що задаються експертом (належність регіону, близькість до тих чи інших транспортних комунікацій і т.д.);

3) формування картографічних діаграм (картограм). У сукупності з розвинутою системою графічного аналізу інформації цей інструмент відкриває користувачам нову форму подання даних, істотно полегшуючи їх оброблення й аналіз;

4) географічна адресна «прив'язка» об'єктів. Цей механізм дає змогу зіставити адресну інформацію об'єкта з «адресним простором» цифрової карти чи плану. Таким чином, виявляються приховані просторові зв'язки між об'єктами досліджуваної системи.

Природно, розробники «Ай-Ту» не могли не використовувати ці очевидні переваги, що істотно розширюють можливості ТІАД. Analyst's Workstation дозволяє проводити «геоприв'язку» об'єктів і формувати картографічні звіти в середовищі Mapinfo (програмний продукт ГІС одного з провідних світових розробників).

Основна проблема при інтеграції з ГІС-пакетами — це їх «закритість». Звичайно повнофункціональна ГІС являє собою досить складний програмний продукт, що містить багато можливостей, призначена для використання фахівцями в галузі картографії і потребує від користувача спеціальної підготовки. Безумовно, подібні системи містять інтерфейсні засоби для імпорту й експорту даних чи навіть мають вбудовані засоби розробки. Але тимчасові і фінансові витрати на інтеграцію залишаються дуже великими і, як правило, не забезпечують повної двосторонньої взаємодії різних частин програмного комплексу.

Велику увагу цій проблемі останнім часом приділяють розробники в сфері ГІС-технологій. Це зумовлено зростаючою потребою у створенні сучасних бізнес-додатків, що включають різноманітні засоби накопичення, аналізу й подання інформації.

Як приклад успішного вирішення, що дозволяє мінімізувати витрати на включення програмних модулів ГІС при проектуванні і реалізації бізнес-додатків, можна навести розробку компанії «ЕРМА ГЕО СОФТ» — бібліотеку геокомпонентів «ГЕО Модуль». Ця технологія дозволяє фахівцям «ЕРМА ГЕО СОФТ» у стислий термін розробляти спеціалізовані бізнес-додатки з урахуванням індивідуальних запитів кожного замовника і легко інтегрувати свої програмні рішення в інформаційну інфраструктуру багатьох компаній.

Незважаючи на широке використання мультимедіа, текст залишається одним з основних видів інформації у більшості електронних сховищ. Розроблення ефективних підходів до оброблення текстів з метою фільтрації, формування змістовного портрету, навігації по базі текстів є одним з найбільш актуальних напрямків сучасних інформаційних технологій. У статті подано спосіб автоматичного нелінгвістичного аналізу

неструктурованої текстової інформації, реалізований на основі нейромережових алгоритмів.

Існуючі підходи до аналізу текстів можна розбити на два класи. До першого класу належать прості, швидкі, незалежні від мови і предметної галузі, але грубі механізми аналізу; найчастіше це підходи, що використовують статистичні методи. Другий клас формують досить витончені, такі, що дають добрий результат, але порівняно повільні підходи, які залежать від мови і предметної галузі; звичайно вони засновані на лінгвістичних методах. Ефективним можна вважати такий підхід, який об'єднав би в собі швидкість і незалежність від мови алгоритмів першого класу з високою якістю оброблення другого.

Пропонований підхід до аналізу текстової інформації реалізовано на основі однорідного нейромережового (а тому статистичного) оброблення інформації, має достатню швидкодію і не залежить від мови і предметної галузі, але при цьому на відміну від більшості алгоритмів оброблення текстів, реалізованих на основі статистичного підходу, дає добрі результати. Так, за даними відомчої експертизи, представлений підхід, реалізований у системі TextAnalyst, визнано кращим у реалізації однієї з важливих функцій оброблення текстів — побудови рефератів — порівняно з підходом, реалізованим, наприклад, на основі лінгвістичних алгоритмів норвезької компанії CognIT¹.

Можливості технології. В існуючих системах оброблення інформації звичайно реалізуються такі основні можливості оброблення текстової інформації: виділення з тексту ключових слів, побудова реферату, формування гіпертекстової структури, ефективна навігація по тексту, порівняння (класифікація) текстів, таксономія множини текстів на рубрики, ефективне подання інформації користувачеві². Пропонований підхід дозволяє реалізувати всі ці функції.

У результаті аналізу тексту з нього автоматично витягується індекс у вигляді мережі основних понять та їх зв'язків з ваговими характеристиками. Як змістовний портрет тексту розглядається не просто список ключових слів, а мережа понять — безліч ключових слів чи стійких словосполучень, пов'язаних між собою. Кожне поняття має деяку вагу, що відображує значущість цього поняття в тексті. Зв'язок між поняттями теж має вагу. Використання зв'язків дозволяє більш точно зважувати поняття тексту.

Структурний портрет тексту (змістовна мережа) може бути побудований як для одного тексту, так і для будь-якої безлічі текстів. Загальна

¹ Див.: Landau D., Feldman R., Aumann Y., Fresko M., Lindell Y., Lipshtat O., Zamir O. TextVis: An Integrated Visual Environment for Text Mining. In Proc. of the 2nd European Symp. on Principles of Data Mining and Knowledge Discovery (PKDD '98), // Lecture Notes in Artif. Intell. 1510, – P. 5664. – Springer-Verlag, 1998.

² Див.: Dan Sullivan. Document Warehousing and Textmining. – NY, 2001.

мережа понять, побудована у великій кількості текстів, що належать до однієї предметної галузі-рубрики, використовується для навігації по рубриці, а також для порівняння з мережами вхідних текстів при їх класифікації. Мінімальний деревоподібний підграф семантичної мережі являє собою тематичне дерево, що, як і семантична мережа, описує зміст тексту, а також дозволяє здійснювати навігацію по тексту. Тематичне дерево більше схоже на зміст тексту. Вихідний текст (безліч текстів) разом з їх семантичною мережею являє собою гіпертекстову структуру і є одночасно сховищем текстів і базою знань. Семантична мережа може бути використана для автоматичного розбивання безлічі текстів на підмножини — таксономії. При цьому групи текстів формуються з урахуванням закладеної в текстах тематичної структури.

Порівняння семантичної мережі вхідного тексту із семантичними мережами рубрик (класифікація) дозволяє зробити висновок про належність тексту до тематики однієї чи декількох рубрик. У результаті класифікації здійснюється зв'язування вхідного тексту з однією чи декількома рубриками, зумовленими користувачем, чи з рубриками, одержаними при таксономії.

Семантична мережа формує метричний простір, в якому текст і будь-яке його речення є векторами. Вибираючи найбільш істотні вектори в цьому просторі (речення), можна автоматично побудувати реферат. Використовуючи тільки частину мережі, що описує деяку тему, можна побудувати реферат і для заданої теми тексту (тематичний реферат) у вигляді послідовності найбільш значущих для заданої теми речень тексту. Вибір підмережі, пов'язаної з деякими поняттями чи групою понять, дає змогу здійснити змістовний підбір текстів (асоціативний пошук) на задану тему.

Аналіз динаміки тематичних таксонів (рубрик), сформованих автоматично, дозволяє визначати появу нової події чи теми в потоці інформації (тематичну динаміку).

Нейромережева основа. В основі оброблення текстової інформації лежить нейромережева технологія¹, єдина для оброблення інформації різних модальностей: текстової, акустичної², візуальної³. При обробленні інформації різних модальностей змінюється тільки спосіб виділення

¹ Див.: Харламов А. А. Автоматический структурный анализ текстов // Открытые системы. – 2002. – № 10.

² Див.: Юрченко А. М. К вопросу определения терминологии и базовых понятий положения идентификации в теории судебной экспертизы для решения задач предупреждения и раскрытия экономических преступлений // Мат. міжнар. наук.-практ. сем. «Сучасні технології у судовій акустиці» (Україна, Київ, 17–18 жовтня 2002 р.) – К., 2003. – С. 46–56.

³ Див.: Hopfield J. J. Neural networks and physical systems with emergent collective computational abilities // Proc. Natl. Acad. Sci. 79, 1982. – P. 2554–2558.

первинних ознак, а при обробленні текстової інформації додатково вводиться зайвий етап оброблення — перенорміровування ваги слів.

Нейромережева технологія оброблення інформації базується на двох фундаментальних посилках. Інформація подається у вигляді одновірної послідовності символів. Починаються спеціальні зусилля для приведення її до такого представлення. Наприклад, для візуальної інформації виділяються крапки зображення, що несуть найбільші зведення, потім виробляється сканування їх послідовності у визначеному порядку, що приводить до формування деякої інформаційної послідовності. Для мови подібні перетворення є простішими, оскільки вона споконвічно одновірна; те ж і з текстами.

Далі ця одновірна інформаційна послідовність відображається в багатомірний сигнальний простір за допомогою вікна, довжиною в n символів таким чином, що кожні n символів є координатами точки в цьому просторі. А всієї послідовності відповідає траєкторія — послідовність точок у цьому просторі. Таке відображення дозволяє відновлювати внутрішню структуру вхідної інформації, якщо така там є. Річ в тім, що це відображення асоціативні за своєю природою. Як тільки вікно заповнене, воно відразу адресує нас до відповідної точки простору. Допустимо, вхідна послідовність має повторювані фрагменти, наприклад слова. Тоді з появою в послідовності слова, яке вже раніше зустрічалося, траєкторія циклічно проходить по своєму фрагменту, що відповідає цьому слову. Якщо є механізм пам'яті, котрий фіксує число проходжень траєкторії через задану точку простору, а також граничне перетворення (тобто перетворення виду $a = 1$, якщо $b > h$, і $a = 0$, якщо $b < h$, де h — поріг), ми здатні усунути з розгляду точки траєкторії, що зустрічаються нечасто, і зберегти точки, що зустрічаються часто. Після такого відсіювання в сигнальному просторі залишається лише інформація, що стосується повторюваних фрагментів послідовності — слів словника заданого рівня.

Після формування такого словника його можна використовувати для фільтрації старої інформації в потоці повідомлень. Такий фільтр пропускає на вихід тільки нову інформацію, яка формує нову послідовність. Вона схожа на стару, з тією відмінністю, що місця в ній, які відповідали старій інформації, замінюються нулями. Ця послідовність зв'язків слів оброблюється на наступному рівні і формує словник уже цього рівня. І так далі. При цьому зв'язки слів словника нижнього рівня враховуються в словах словника наступного рівня.

У разі оброблення текстів алгоритм трохи спрощується; залишаються тільки два рівні аналізу. При цьому зберігається основна мета — сформу-

вати словник слів і виявити зв'язок між ними, тому на першому кроці формується словник слів. Але на другому кроці замість формування послідовності абrevіатур шляхом фільтрації слів словника, сформованого на першому рівні, аналізується попарна зустрічальність слів у реченнях.

Таким чином, виявляються зв'язки між словами. Одержана інформація являє собою частотний портрет тексту, що може бути візуалізований у вигляді мережі слів словника та їх зв'язків. Причому і слова словника, й їх зв'язок одержують у результаті аналізу деяких частотних характеристик.

Для оброблення текстів на відміну від мови і відео вводиться ще одна операція — перенорміровування числових характеристик слів. Для цього використовується ітеративний алгоритм роботи нейронних мереж Хопфилда¹. На кожній ітерації слово мережі збільшує свою вагу, якщо воно виявляється пов'язаним з багатьма іншими словами з великою вагою, а інші слова її рівномірно втрачають. Так, якщо слово зустрілося двічі — у заголовку й анотації, воно є ключовим поняттям і може придбати дуже велику вагу.

Можливі додатки. Рішення для пошуку інформації в Web. Дана технологія дозволяє за допомогою пошукових програмних агентів автоматично підібрати інформацію із заданої теми і створити базу знань. Web використовується як глобальне сховище знань. Технологія пошукових програмних агентів дає змогу здійснити тематично спрямоване збирання інформації в Internet і формування бази знань. Побудова бази знань починається з теми, що може бути задана одним-двома словами. При пошуку вона обростає значеннями доповненнями (вони виходять внаслідок розширення асоціаціями на асоціативній мережі), по яких пошукові програмні агенти теж збирають інформацію.

Персоніфікація процесу збирання й аналізу інформації. Сформована персональна база знань користувача стає своєрідним фільтром, що дозволяє підвищити точність виконання таких операцій, як пошук у Internet і моніторинг нової інформації. При цьому визначаються і враховуються переваги користувача. Під час пошуку інформації в Internet спеціальний програмний агент переглядає зміст сторінок, на які заходить користувач, зчитує з них інформацію і поміщає її в персональну базу знань користувача, в якій на верхніх рівнях відновлюються його основні інтереси і переваги. Використовуючи персональну базу знань, програмні агенти можуть автоматично збирати нову інформацію, що відповідає інтересам користувача. Персоніфікована фільтрація дозволяє автоматично вибрати тільки ту інформацію, що може виявитися цікавою.

Повнотекстові бібліотеки. Асоціативна мережа може бути використана поряд зі стандартними рубриками для навігації по базі текстів.

¹ Див.: Hopfield J. J. Вказ. праця.

Фільтрація текстів. Така система може використовуватися з метою «не випускати» чи «не впускати» повідомлення, наприклад, для недопущення витоку конфіденційної інформації. Вона ж може використовуватися для класифікації текстів по рубриках (адресатам).

Оболонка для електронних книг. Гіпертекстове подання є зручним для створення електронних книг.

Висновок. Програма TextAnalyst дозволяє автоматично сформувати змістовний портрет тексту у вигляді асоціативної мережі основних понять з їх зв'язками, позначеними їх числовими характеристиками. Семантична мережа разом з вихідним текстом являють собою гіпертекстову структуру, на якій семантична мережа відіграє роль ефективного засобу навігації. Розбивання мережі безлічі текстів на підмережі робить автоматичну таксономію безлічі текстів по рубриках. Порівняння семантичних мереж текстів дозволяє виявити ступінь кореляції між текстами, а порівняння семантичних мереж вхідного тексту і рубрик — віднести текст до однієї з рубрик. Елементарно будується реферат тексту. Так само просто будується тематичний реферат і здійснюється змістовний пошук на безлічі текстів. Оскільки нейромережева технологія реалізує специфічний статистичний аналіз, оброблення тексту не залежить від предметної галузі і мови, виконується швидко і не вимагає дискового простору. В даний час TextAnalyst підтримує сім мов, включаючи китайську.

Надійшла до редколегії 15.09.05

О. Наден, доцент Східноукраїнського національного університету

Щодо методологічних засад розуміння кримінального права (у зв'язку з дослідженнями проблем кримінально-правового регулювання)

Існування суспільства зумовлює необхідність регулювання його внутрішніх процесів — відносин між його учасниками. І чим більш різноманітними є такі відносини, тим різноманітнішим є й інструментарій для їх регулювання. Більше того, в різні історичні періоди на перший план у регулюванні суспільних відносин виходять різні регулятивні системи. Історія людства знає часи, коли головну роль у регулю-

ванні відносин у суспільстві відіграли мораль, релігія, позаправове свавілля. Сьогодні в ньому провідне місце посідає право як певна сукупність специфічних соціальних норм. Саме тому особливості регулятивного впливу права на суспільні відносини становлять інтерес для всіх соціальних наук, передусім правознавства.

Радянське правознавство однією з основних своїх методологічних засад мало твердження про єдність системи радянського права, наслідком чого були спроби виробити єдине поняття права, єдине вчення про механізм правового регулювання і його складові системи: норми права, правовідносини, акти застосування права, єдиних класифікаційних критеріїв для різних правових явищ тощо. Ці намагання мали певний успіх й їх результати спрацьовували тоді, коли дійсно «для нас нічого приватного не було», коли все право було опублічене. Сьогодні ж, коли наша правознавча наука повертається до своїх методологічних першоджерел, що є загальноприйнятими принаймні в рамках романо-германської правової сім'ї, ми все частіше стикаємося з тим, що розроблені в радянський період категорії, які мали б бути універсальними для всієї системи права, по-різному проявляються в окремих його галузях, а подеколи не спрацьовують взагалі. І головний обсяг таких неузгодженостей спостерігається там, де ми починаємо виділяти в праві два основні напрямки — право приватне і право публічне.

Саме тому однією з нагальних проблем сучасного правознавства вбачається вироблення основних методологічних засад розуміння права, які узгоджувалися б з особливостями приватноправової чи публічно-правової природи його норм. Лише на цьому підґрунті і можуть в подальшому розглядатися інші, більш-менш обсяжні питання правового регулювання в тій чи іншій сфері. При цьому не можна не відзначити, що в науці кримінального права розуміння цього вже складається, свідченням чого є дослідження, що проводяться як українськими (Ю. Баулін, М. Мельник, В. Навроцький та ін.), так і російськими (Є. Благгов, В. Дуюнов, В. Коняхін, Н. Лопашенко, Г. Петрова та ін.) криміналістами. Їхніми зусиллями розробляються такі методологічні проблеми науки, як публічно-правовий статус кримінального права, зміст і значення кримінально-правового регулювання, норми кримінального права та кримінальні правовідносини тощо. Разом з тим цілком зрозумілою є ситуація, за якої кожен дослідник знаходиться в колі своїх «масштабів» та «координат», тим самим прагнучи до застосування методологічних розробок щодо своїх прикладних питань. У зв'язку з цим схожі концептуальні засади різні автори часто бачать по-різному, а деякі з них не розглядаються взагалі. Проте визначення в концептуальних характеристиках кримінально-