

Федеральное государственное бюджетное учреждение науки
ЦЕНТРАЛЬНЫЙ ЭКОНОМИКО-МАТЕМАТИЧЕСКИЙ ИНСТИТУТ РАН
CENTRAL ECONOMICS AND MATHEMATICS INSTITUTE RAS

РОССИЙСКАЯ
АКАДЕМИЯ НАУК

RUSSIAN
ACADEMY OF SCIENCES

Л.Е. Варшавский

**СОЦИАЛЬНО-ЭКОНОМИЧЕСКИЕ
ПРОБЛЕМЫ РАЗВИТИЯ
ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫХ
ТЕХНОЛОГИЙ (ИКТ)**

Москва
2022

Варшавский Л.Е. Социально-экономические проблемы развития информационно-коммуникационных технологий (ИКТ) [Текст]: монография. М.: ЦЭМИ РАН, 2022. – 158 с. (Рус.)

Исследуются проблемы, связанные с развитием рынков компонентов и аппаратных средств информационно-коммуникационных технологий (ИКТ). Проводится анализ состояния и развития мирового рынка полупроводниковой продукции, а также рынка микропроцессоров. Рассматриваются экономико-математические модели развития крупных компаний-производителей интегральных схем и микропроцессоров. Исследуются сценарии среднесрочного прогнозирования динамики ключевых показателей рынка микропроцессоров, а также рынка мобильной связи, построенные на основе разработанных игровых моделей. Проводится технико-экономический анализ тенденций развития суперкомпьютеров, а также анализ проблем, связанных с развитием индустрии квантовых вычислений.

Анализируются социально-экономические проблемы, связанные с расширением использования ИКТ. Значительное внимание уделяется исследованию факторов риска, связанных с развитием производственной базы ИКТ-компаний и предприятий, производящих аппаратные средства ИКТ, а также факторов риска для общества. Проводится анализ национальных программ развития технологий искусственного интеллекта (ИИ), а также прогнозных сценариев оценки масштабов и социально-экономических последствий применения методов ИИ. Рассматриваются социально-экономические проблемы, связанные с использованием ИКТ в «умных» городах.

Рассматриваемый в монографии материал ориентирован на использование в практике планирования, прогнозирования и государственного регулирования экономистами, инженерами, а также специалистами в области естественных наук. Книга может быть также рекомендована студентам старших курсов и аспирантам экономических, гуманитарных и ряда инженерных специальностей (радиоэлектроника, вычислительная техника, системы автоматического управления и др.).

Ключевые слова: ИКТ, социально-экономические проблемы, интегральные схемы, рынки, экономико-математические модели, микропроцессоры, мобильная связь, суперкомпьютеры, квантовые вычисления, факторы риска, «умные» города.

Рецензенты: д.э.н., проф. Е.Р. Орлова (ФИЦ «Информатика и управление»);
д.э.н., проф. Б.А. Ерзнкян (ЦЭМИ РАН).

Varshavsky L.E. Socio-economic problems of the development of information and communication technologies (ICT). [Text]: monograph. M.: CEMI RAS, 2022. – 158 p. (Rus.)

The problems associated with the development of markets for components and hardware of information and communication technologies (ICT) are studied. An analysis of the state of the world market for semiconductor products, and of the development for computer system components markets is being carried out. Economic and mathematical models of the development of large manufacturers of integrated circuits and of microprocessors are considered. Scenarios for medium-term forecasting of the dynamics of key indicators of the microprocessor market, as well as of the mobile communications market for the medium term, built using dynamic game models, are studied.

A technical and economic analysis of the evolution of high-performance supercomputers and of the problems associated with the development of the quantum computing industry are carried out.

The socio-economic problems associated with the expansion of the use of ICT are analyzed. Considerable attention is paid to the study of risk factors associated with the development of the companies and enterprises that produce ICT hardware, as well as risk factors for society. An analysis is being made of national programs for the development of artificial intelligence (AI) technologies, as well as predictive scenarios for assessing the scale and socio-economic consequences of using AI. Social and economic problems are considered, related to the use of ICT in smart cities. The material considered in the monograph is focused on use in the practice of planning, forecasting and state regulation by economists, engineers, as well as specialists in the field of natural sciences. The book can also be recommended to senior students and graduate students in economics, the humanities and a number of engineering specialties (radio electronics, computer technology, control systems, etc.).

Keywords: ICT, socio-economic problems, integrated circuits, markets, economic and mathematical models, microprocessors, mobile communications supercomputers, quantum computing industry, risk factors, smart cities.

ОГЛАВЛЕНИЕ

| | |
|---|----|
| ВВЕДЕНИЕ..... | 5 |
| <i>Список литературы к введению.....</i> | 7 |
| РАЗДЕЛ 1. АНАЛИЗ СОСТОЯНИЯ И ПЕРСПЕКТИВ РАЗВИТИЯ РЫНКОВ КОМПОНЕНТОВ ИКТ И РЫНКА МОБИЛЬНОЙ СВЯЗИ | 9 |
| Глава 1. Современные тенденции развития полупроводниковой промышленности..9 | |
| 1.1. Введение..... | 9 |
| 1.2. Особенности производственного цикла и глобализации полупроводниковой промышленности | 12 |
| 1.3. Роль США и Китая в глобализации полупроводниковой промышленности | 16 |
| 1.4. Некоторые перспективные направления развития полупроводниковой отрасли | 18 |
| 1.5. Развитие рынка полупроводников в перспективе | 19 |
| 1.6. Моделирование динамики показателей рынка высокотехнологичной продукции с догоняющим участником (на примере фабричного производства чипов) | 23 |
| 1.7. Заключение..... | 33 |
| Список литературы к главе 1..... | 34 |
| Глава 2. Моделирование динамики ключевых показателей рынка микропроцессоров | 36 |
| 2.1. Введение..... | 36 |
| 2.2. Основные тенденции развития элементной базы НРС (на примере микропроцессоров)..... | 38 |
| 2.3. Экономико-математический анализ и моделирование развития ведущих компаний микроэлектронной промышленности..... | 45 |
| 2.4. Прогнозирование динамики показателей рынка микропроцессоров x86 на среднесрочный период | 47 |
| 2.5. Исследование перспективной структуры рынка микропроцессоров для серверов (с архитектурой x86 и ARM) | 54 |
| 2.6. Возможный подход к формированию рыночных структур в микроэлектронике..... | 60 |
| 2.7. Заключение..... | 63 |
| Список литературы к главе 2..... | 63 |
| Глава 3. Исследование перспективной динамики показателей развития рынка мобильной связи | 66 |
| 3.1. Введение..... | 66 |
| 3.2. Структура модели | 68 |
| 3.3. Прогнозные сценарии | 70 |
| 3.4. Некоторые препятствия на пути широкомасштабного развертывания новых технологий мобильной связи в США..... | 75 |
| 3.5. Заключение..... | 75 |
| Список литературы к главе 3..... | 76 |

| | |
|---|-----|
| РАЗДЕЛ 2. ТЕХНИКО-ЭКОНОМИЧЕСКИЙ АНАЛИЗ ТЕНДЕНЦИЙ РАЗВИТИЯ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ..... | 77 |
| Глава 4. Исследование динамики ключевых показателей высокопроизводительных суперкомпьютеров..... | 77 |
| 4.1. Введение..... | 77 |
| 4.2. Анализ ключевых технико-экономических показателей оборудования для суперкомпьютеров..... | 78 |
| 4.4. Проблема создания суперкомпьютера экзафлопсной производительности..... | 88 |
| 4.5. Заключение..... | 92 |
| Список литературы к главе 4..... | 92 |
| Глава 5. Проблемы развития индустрии квантовых вычислений..... | 94 |
| 5.1. Введение..... | 94 |
| 5.2. Особенности квантовых компьютеров..... | 94 |
| 5.3. Некоторые подходы к созданию квантовых компьютеров..... | 96 |
| 5.4. Гонка за лидерство в области квантовых технологий..... | 100 |
| 5.5. Принципиальные проблемы, требующие решения для практического использования квантовых вычислений..... | 102 |
| 5.6. Экосистема индустрии квантовых вычислений..... | 104 |
| 5.7. Заключение..... | 105 |
| Список литературы к главе 5..... | 106 |
| РАЗДЕЛ 3. АНАЛИЗ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ПРОБЛЕМ, СВЯЗАННЫХ С ИКТ | 108 |
| Глава 6. Исследование факторов риска ИКТ на уровне экономических агентов ... | 108 |
| 6.1. Введение..... | 108 |
| 6.2. Экономический анализ факторов риска производителей аппаратных средств ИКТ..... | 109 |
| 6.3. Анализ факторов риска потребителей ИКТ-продуктов и услуг..... | 115 |
| 6.4. Моделирование динамики экономических показателей рынков при наличии киберугроз (на примере олигополии)..... | 122 |
| 6.5. Заключение..... | 130 |
| Список литературы к главе 6..... | 130 |
| Глава 7. Социально-экономические факторы риска ИКТ для общества | 132 |
| 7.1. Введение..... | 132 |
| 7.2. Влияние технологий искусственного интеллекта на развитие общества..... | 133 |
| 7.3. Влияние ИИ на рынок труда..... | 138 |
| 7.4. Социальные факторы риска для общества..... | 141 |
| 7.5. Социальные проблемы, связанные с распространением ИКТ в России..... | 146 |
| 7.6. Проблемы и риски, связанные с созданием «умных» городов..... | 147 |
| 7.7. Заключение..... | 155 |
| Список литературы к главе 7..... | 156 |

ВВЕДЕНИЕ

В настоящее время информационно-коммуникационные технологии (ИКТ) превратились в важную отрасль, определяющую будущее развитие экономики. По оценке компании Gartner в 2021 г. объем мирового рынка ИКТ составлял 4,26 трлн долл., в том числе объем только информационных технологий (ИТ) – 2,82 трлн долл., а коммуникационных – 1,44 трлн долл. В структуре расходов на информационные технологии около 64% составляют расходы на информационные услуги, а 36% – расходы на аппаратные средства (devices) и на системы центров обработки данных (ЦОД)ⁱ.

Развитие и использование информационно-коммуникационных технологий привело к революционным преобразованиям в экономике. Особенно наглядно положительная роль ИКТ была продемонстрирована в разгар пандемии COVID-19 в 2020–2021 гг., когда, несмотря на значительные потери, за счет использования этих технологий удалось избежать полного хозяйственно-экономического хаоса. Однако побочным эффектом распространения этих технологий стало появление все новых киберугроз и рисков, которые в перспективе начнут сдерживать внедрение ИКТ.

Недоучет потенциальных ИКТ-рисков может в будущем привести к серьезным последствиям для общества. Так, по оценкам Cybersecurity Ventures, глобальные потери в результате киберпреступлений увеличатся с 3 трлн долл в 2015 г. до 10,5 трлн долл в 2025 г.ⁱⁱ

Цель настоящей работы состоит в анализе социально-экономических проблем, связанных с развитием рынков компонентов и аппаратных средств информационно-коммуникационных технологий (ИКТ), а также факторов риска, учет и подготовка к которым могли бы способствовать смягчению неблагоприятных для общества последствий от слабо контролируемого в настоящее время развития ИКТ. Эти проблемы актуальны для нашей страны, в которой реализуется принятая в 2017 г. Программа «Цифровая экономика Российской Федерации». Успешное выполнение этой Программы в условиях все расширяющихся санкций невозможно без усиления отечественной производственной базы в вышеуказанных областях.

Монография состоит из трех разделов. В первом разделе, включающем три главы, рассматриваются состояние и перспективы развития рынка компонентов ИКТ (на примере рынков интегральных схем и микропроцессоров), а также рынка мобильной связи.

ⁱ URL: <https://www.gartner.com/en/newsroom/press-releases/2022-04-06-gartner-forecasts-worldwide-it-spending-to-reach-4-point-four-trillion-in-2022> (дата доступа 29.04.2022).

ⁱⁱ URL: <https://cybersecurityventures.com/hackerpocalypse-cybercrime-report-2016/> (дата доступа 20.04.2022).

В первой главе проводится краткий анализ состояния мирового рынка полупроводниковой продукции в последние годы. Рассматриваются особенности производственного цикла в отрасли. Дается описание основных характерных черт глобализации в производстве полупроводников. Рассматриваются разработанные с помощью экономико-математической модели сценарии развития рынка интегральных схем, а также возможные потери, связанные с серьезным нарушением цепи поставок в среднесрочной перспективе. С использованием теории динамических игр исследуются динамика структуры олигополистических рынков в условиях вхождения на них новых догоняющих участников, руководствующимися неэкономическими критериями (на примере рынка производителей пластин с микросхемами). Материал первой главы основан на статьях автора [1, 2].

Во второй главе проводится технико-экономический анализ развития рынка микропроцессоров для вычислительных систем. Рассматриваются экономико-математические модели развития крупных компаний-производителей микропроцессоров и микросхем. С использованием игровой модели исследуются сценарии среднесрочного прогнозирования динамики ключевых показателей рынка микропроцессоров архитектуры x86, а также рынка микропроцессоров для серверов. Материал второй главы основан на статьях автора [3–5].

В третьей главе проводится анализ показателей развития олигополистического рынка мобильной связи (на примере США) на средне- и долгосрочную перспективу. С использованием методов теории динамических игр рассматриваются последствия возможного объединения компаний-провайдеров T-Mobile и Sprint. Путем компьютерного моделирования исследованы условия, при которых может быть повышена рыночная доля объединенной компании. Материал третьей главы основан на статье автора [6].

Во втором разделе проводится технико-экономический анализ тенденций развития суперкомпьютеров и квантовых компьютеров.

В четвертой главе раздела исследуются тенденции изменения показателей развития высокопроизводительных суперкомпьютеров. Рассматриваются разработанные автором модели динамики распространения высокопроизводительных суперкомпьютеров, структуры вычислительных мощностей, а также «гонки» производительности суперкомпьютеров между отдельными странами. Приводятся прогнозные оценки ключевых показателей суперкомпьютеров на кратко- и среднесрочный период. Материал четвертой главы основан на статьях автора [7, 8].

В пятой главе исследуются проблемы, связанные с развитием индустрии квантовых вычислений. Кратко рассматриваются подходы к созданию квантовых компьютеров, а также принципиальные проблемы, затрудняющие переход к

практическому их использованию в ближайшей перспективе. Исследуется экономическая среда, в которой происходит эволюция индустрии квантовых вычислений. Материал пятой главы основан на статье автора [9].

Третий раздел, состоящий из двух глав, посвящен анализу социально-экономических проблем, связанных с расширением использования ИКТ, учет и решение которых могли бы способствовать смягчению неблагоприятных для развития экономических агентов последствий.

В шестой главе значительное внимание уделяется анализу факторов риска, связанных с развитием производственной базы ИКТ-компаний и предприятий, производящих аппаратные средства ИКТ. Рассматриваются экономические факторы риска использования ИКТ для экономических объектов, а также игровой подход к исследованию динамики показателей олигополистических рынков высокотехнологичной продукции, в которых производственная инфраструктура участников подвергаются кибератакам. Рассматриваются результаты расчетов показателей условного олигополистического рынка, характерного для высокотехнологических производств, при разных сценариях восприятия интенсивности киберугроз олигополистами. Материал шестой главы основан на статьях автора [10, 11].

В седьмой главе исследуются социально-экономические факторы риска ИКТ для общественного развития. Рассматриваются тенденции развития технологии искусственного интеллекта (ИИ). Проводится анализ национальных программ развития технологий ИИ, а также прогнозных сценариев оценки масштабов и социально-экономических последствий применения методов ИИ. Рассматриваются риски, связанные с продукцией и услугами отрасли ИКТ. Рассматриваются также социально-экономические проблемы, связанные с использованием ИКТ в «умных» городах. Материал седьмой главы основан на статьях автора [12, 13].

Автор выражает глубокую благодарность рецензентам, д.э.н., проф. Б.А. Ерзнкяну (ЦЭМИ РАН) и д.э.н., проф. Е.Р. Орловой (ФИЦ «Информатика и управление») за ценные замечания и предложения, позволившие улучшить содержание настоящей монографии.

Список литературы к введению

1. Варшавский Л.Е. Современные тенденции развития полупроводниковой промышленности // Концепции. 2021. № 1 (40). С. 44–50.
2. Варшавский Л.Е. Исследование эволюции рынка высокотехнологичной продукции с догоняющим участником // Теория и практика институциональных преобразований в России / Ерзнкян Б.Г. (ред.). Вып. 53. М.: ЦЭМИ РАН, 2022. С. 61–70.

3. Варшавский Л.Е. Моделирование динамики ключевых показателей рынков компонентов высокопроизводительных вычислительных систем // Труды Института системного анализа Российской академии наук. 2017. Т. 67. № 1. С. 12–27.
4. Варшавский Л.Е. Анализ инвестиционных процессов в ведущих компаниях микроэлектронной промышленности // Концепции. 2016. № 1 (35). С. 17–28.
5. Варшавский Л.Е. Экономико-математические методы анализа динамики рыночных процессов и формирования конкурентной среды в микроэлектронной промышленности // Национальные интересы: приоритеты и безопасность. 2016. № 336. С. 26–44.
6. Варшавский Л.Е. Исследование показателей развития рынка мобильной связи // Концепции. 2019. № 1 (38). С. 36–45.
7. Варшавский Л.Е. Экономико-математическое моделирование тенденций изменения ключевых показателей высокопроизводительных суперкомпьютеров // Труды Института системного анализа Российской академии наук. 2017. Т. 67. № 3. С. 64–74.
8. Варшавский Л.Е. Анализ и моделирование динамики ключевых показателей высокопроизводительных суперкомпьютеров // Концепции. 2015. № 2 (34).
9. Варшавский Л.Е. Текущие проблемы развития квантовых вычислений // Теория и практика институциональных преобразований в России: сб. научных трудов / Ерзнкян Б.Г. (ред.). Вып. 47. М.: ЦЭМИ РАН, 2019. С. 33–45.
10. Варшавский Л.Е. Анализ факторов риска, связанных с информационно-коммуникационными технологиями // Концепции. 2017. № 1 (36). С. 37–47; пп. 08-09.
11. Варшавский Л.Е. Моделирование развития олигополистических рынков при наличии киберугроз // Труды Института системного анализа Российской академии наук. 2020. Т. 70. Вып. 2. С. 21–29.
12. Варшавский Л.Е. Анализ социально-экономических проблем и вызовов, связанных с ИКТ // Труды Института системного анализа Российской академии наук. 2019. Т. 69. Вып. 1. С. 3–16.
13. Варшавский Л.Е. Концепция «умного города» и риски информационно-коммуникационных технологий (ИКТ) // Теория и практика институциональных преобразований в России: сб. научных трудов / Ерзнкян Б.Г. (ред.). Вып. 50. М.: ЦЭМИ РАН, 2020. С. 116–127.

РАЗДЕЛ 1.

АНАЛИЗ СОСТОЯНИЯ И ПЕРСПЕКТИВ РАЗВИТИЯ РЫНКОВ КОМПОНЕНТОВ ИКТ И РЫНКА МОБИЛЬНОЙ СВЯЗИ

ГЛАВА 1.

СОВРЕМЕННЫЕ ТЕНДЕНЦИИ РАЗВИТИЯ ПОЛУПРОВОДНИКОВОЙ ПРОМЫШЛЕННОСТИⁱⁱⁱ

1.1. Введение

Полупроводниковая промышленность, возникшая во второй половине XX в., является главным драйвером развития ИКТ и одним из драйверов развития экономики и общественного развития в целом. В истории возникновения и развития полупроводников условно выделяют четыре этапа^{iv}:

- 1) создание и использование вакуумных ламп (1901–1945 гг.);
- 2) создание транзисторов (1945–1950-е гг.);
- 3) создание интегральных схем (1960-е гг.);
- 4) создание и микропроцессоров (1970–1990-е гг.).

Возрастающий спрос на полупроводниковую промышленность обусловлен непрерывным появлением инновационных технологий. Так, в 1997–2007 гг. в связи с быстрым распространением персональных компьютеров (ПК), повысился спрос на центральные процессоры (ЦП) и микросхемы памяти. Широкое проникновение Интернета привело к увеличению объемов оборудования Ethernet и сетевых процессоров. Появление смартфонов iPhone в 2007 г. стимулировало спрос на мобильные процессоры, в то время как внедрение облачных вычислений стимулировало рост серверных процессоров и хранилищ. В настоящее время считается, что новый десятилетний цикл роста полупроводникового сектора будет связан с форсированным переходом к искусственному интеллекту (ИИ)^v.

С момента изобретения интегральной схемы в 1958 г. количество транзисторов на пластину для логической микросхемы увеличилось примерно в 10 млн раз, что привело к увеличению скорости процессора в 100 тыс. раз и снижению затрат

ⁱⁱⁱ Материал данной главы основан на статьях автора [1, 2].

^{iv} URL: <https://www.wipo.int/publications/ru/details.jsp?id=3998&plang=ES> (дата обращения 25.05.2021).

^v URL: <https://www.pwccn.com/en/industries/telecommunications-media-and-technology/publications/global-tmt-semiconductor-report-2019.html> (дата обращения 25.05.2021).

более чем на 45% на каждую пластину. Это позволило резко увеличить вычислительные мощности устройств при одновременном уменьшении их габаритов. В результате вычислительная мощность современных смартфонов превышает мощности мейнфреймов, использовавшиеся НАСА для обеспечения полета космического аппарата Аполлон-11 на Луну в 1969 г. Отмечается также, что современные смартфоны содержат больше памяти для хранения, чем сервер центра обработки данных (ЦОД) в 2010 г. [3].

В связи с быстрым внедрением инноваций в полупроводниковой отрасли, ее влияние на экономику все более усиливается. Так, по данным IC Insights, если в 1980–1989 гг. коэффициент корреляции глобального ВВП и рынка интегральных схем (ИС) (последний составляет доминирующую долю рынка полупроводников), составлял 0,35, то в 2000–2009 гг. он повысился до 0,63, а в 2010–2018 гг. – до 0,87 [4].

Интересно отметить, что в структуре производимых полупроводников и ИС существенно снизилась доля компьютеров, но возросла доля таких отраслей, как связь и автомобилестроение (табл. 1.1). Лидирующую позицию в производстве интегральных схем занимают фирмы США (52% мирового производства), за которыми следуют фирмы Южной Кореи [5].

Таблица 1.1

Структура рынка интегральных схем, % [4]

| Год | Компьютеры | Связь | Потребительская электроника | Автомобилестроение | Другие отрасли промышленности |
|------|------------|-------|-----------------------------|--------------------|-------------------------------|
| 1998 | 55,6 | 18,5 | 13 | 4,7 | 8,2 |
| 2003 | 51,3 | 22 | 12,9 | 6,3 | 7,5 |
| 2008 | 45 | 29,3 | 12,2 | 6,5 | 7 |
| 2013 | 36,7 | 37,2 | 12,5 | 7,1 | 6,5 |
| 2019 | 35,6 | 35,6 | 11,8 | 8,7 | 8,3 |

Общий объем поставок полупроводниковых элементов, включая интегральные схемы, а также оптоэлектронику, датчики, исполнительные механизмы и дискретные (OSD) устройства, по оценкам McClean IC Insights достиг в 2021 г. рекордного уровня 1,3 трлн единиц. При этом совокупный годовой темп роста (CAGR) для полупроводниковых устройств за период с 1978 г., когда было произведено 32,6 млрд единиц, по 2021 г., превысил 9% [6].

Глобальный рынок полупроводников, включая интегральные схемы, а также устройства оптоэлектроникт, датчики / исполнительные механизмы и дискретные устройства (OSD) достиг в 2021 г. 614 млрд долл., в том числе ИС – почти 510 млрд долл., что на 88,5% превышает уровень 2011 г.! В текущем 2022 г., по оценкам IC

Insights, общий объем производства полупроводников увеличится до 681 млрд долл., а ИС – до 565 млрд долл. [6].

Вместе с тем, следует отметить неравномерность темпов роста рынка полупроводников и особенно ИС. Так, основной рост произошел в 2016–2021 гг. (на 71%), после относительно стабилизации в 2011–2015 гг. (рис. 1.1). Особенно заметный рост рынка произошел в 2019–2021 гг., что связано с интенсификацией торговой войны между США и Китаем в 2018–2019 гг., а также с распространением COVID-19.

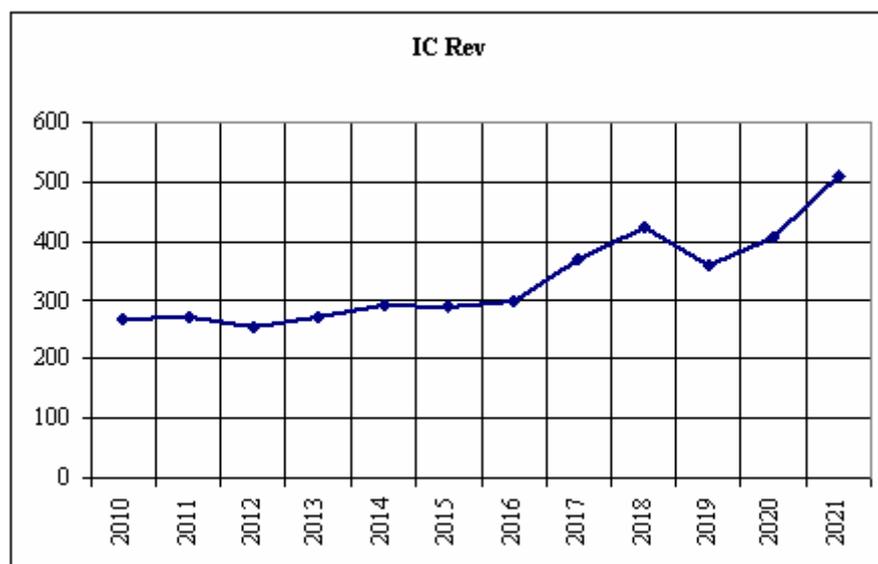


Рис. 1.1. Динамика объемов мирового рынка ИС (IC Rev, млрд долл.)

Более того, из-за распространения COVID-19 и связанных с этим производственных и логистических проблем, в последние годы, начиная с 2020 г., проявляется острый дефицит полупроводников. В значительной степени это обусловлено ростом спроса на устройства для поддержания удаленной работы и обучения, а также ростом облачных вычислений. Из-за дефицита полупроводниковой продукции существенно сократилось производство в автомобильной промышленности, которая становится крупным потребителем полупроводников. Вместе с тем, обострился дефицит полупроводниковых комплектующих для компьютеров, который, по оценкам ряда компаний, может продлиться и до 2023–2025 гг. В связи с этим, существенно выросли цены на компьютеры и ноутбуки (за 2021 г., например, на российском рынке – на 25–30% [7]).

1.2. Особенности производственного цикла и глобализации полупроводниковой промышленности

1.2.1. Особенности производственного цикла

Изготовление полупроводниковой микросхемы – отличается весьма высокой капиталоемкостью и нукоемкостью. При производстве только полупроводниковых пластин может быть до 1400 этапов процесса (в зависимости от сложности процесса). Изготовление готового чипа для заказчика может занять до 26 недель. Так, изготовление готовой полупроводниковой пластины, известное как время цикла, занимает в среднем около 12 недель, но для сложных процессов может потребоваться до 14–20 недель. Для отладки процесса изготовления микросхемы с целью увеличения объемов производства, требуется еще больше времени – около 24 недель [8].

После завершения процесса изготовления микросхем, осуществляется заключительный этап производства, включающий в себя внутреннюю сборку, тестирование и упаковку, на выполнение которого уходит до 6 недель. В целом, время выполнения заказа с момента размещения заказа до получения конечного продукта может занять в общей сложности до 26 недель [8].

Создание такого высокотехнологичного производства требует колоссальных все возрастающих инвестиций в исследования и разработки, и, непосредственно в производство. В настоящее время, как нукоемкость, так и капиталоемкость в отрасли превышают 20% (по отношению к объемам дохода). В связи с колоссальной нуко- и капиталоемкостью отрасли, уже в 1980–1990-е гг. произошло разделение фирм на фирмы-производители пластин с микросхемами (англ. Foundries) и большое число фирм-разработчиков дизайна схем и приборов (англ. Fabless, т.е. не имеющие в своем составе фабрик) [10].

Среди первой группы компаний лидирующая роль принадлежит тайваньской TSMC со своим гигантским комплексом GigagFab. Концепция Foundry-фирмы была успешно реализована на Тайване М.Чангом в ходе создания компании TSMC при поддержке местного правительства [9]. Эта компания, созданная в конце 1980-х гг., превратилась в крупнейшую компанию-производителя пластин и микросхем (Pure-Play Semiconductor Foundry) на основе дизайнов второй группы компаний (fabless), с годовым объемом продаж в настоящее время более 50 млрд долл. На нее приходится почти 60% объема продаж фирм первой группы. За последнее десятилетие повысилась концентрация производства на рынке производителей пластин (Foundries), в результате чего он превратился в чисто олигополистический, с

небольшим числом крупных участников. На этом рынке доля 10 наиболее крупных производителей составляет свыше 90%.

Уровень концентрации на рынке фирм-разработчиков (Fabless), суммарные доходы которых в 2020 г. оценивались почти в 133 млрд долл. – несколько ниже (доля 10 крупнейших фирм составляет примерно 2/3)^{vi}. Наиболее крупными из них являются компании Qualcomm, Broadcom, AMD (с 2009 г., после передачи своих фабрик в компанию Global Foundries).

Среди чисто дизайнерских фирм особая роль принадлежит компании ARM Ltd. которая специализируется на разработке чипов с процессорами на основе собственной ARM-архитектуры. По данным этой компании, в 2015 г. с использованием этих процессоров всего в мире было произведено 15 млрд чипов, причем 45% из них было использовано в мобильных устройствах. В 85% всех производимых смартфонов были установлены энергоэффективные процессоры с архитектурой ARM. Сравнительно небольшие доходы этой компании (в 2015 г. они составляли немногим менее 1,5 млрд долл) формируются за счет лицензий и ройалти на производимую продукцию. [10].

Такое разделение позволило фирмам-разработчикам схем сосредоточиться на стадии исследований и разработок, а также на продаже продукции. За счет аутсорсинга производства им удалось существенно снизить зависимость от выпуска своей продукции на свободных мощностях фабрик крупнейших компаний и, тем самым, риск потери прав на интеллектуальную собственность. Наряду с перечисленными выше группами фирм, к настоящему времени осталось небольшое число гигантских интегрированных компаний (англ. integrated device manufacturer (IDM)), играющих ключевую роль в отрасли, которые сами разрабатывают, производят и продают схемы и приборы. К их числу относятся такие крупнейшие компании, как Intel, Samsung, IBM, STMicroelectronics и др. [10].

Об объемах капитальных вложений в отрасли свидетельствует тот факт, что такие ведущие компании, как Intel, Samsung и TSMC уже продолжительное время ежегодно инвестируют в производство более 10 млрд долл. В последнее время Samsung и TSMC существенно превзошли по инвестициям Intel, который долгое время осуществлял максимальные среди компаний отрасли капитальные вложения. С 2017 г. годовые капитальные вложения Samsung в полупроводниковое производство превышают 20 млрд долл, причем в 2020 г. они составили 28,1 млрд долл., а за период 2017–2020 гг. – 93.2 млрд долл. В 2021 г. TSMC, уверенно освоившая 7- и

^{vi} URL: http://www.electronics-eetimes.com/en/fabless-chip-companies-ranked-by-2013-sales.html?cmp_id=7&news_id=222921061

5-нм процессы (во 2-м квартале 2020 г. на микросхемы этих размеров приходилось 47% продаж компании), планировала увеличить свои капитальные расходы до 25–28 млрд долл. [11]. Более того, с целью устранения возможного дефицита в будущем, TSMC в начале 2021 г. объявила об инвестиционном плане расширения действующих и сооружения новых производственных объектов отрасли на общую сумму в 100 млрд долл., который должен быть реализован в течение трехлетнего периода до 2023 г. В этом плане намечено расширение существующих фабрик, а также строительство новых заводов по производству микросхем в США и других странах [12].

1.2.2. Глобализация полупроводниковой промышленности

В связи со сложностью высокотехнологического производства полупроводниковой продукции, зависящего от длинной цепи поставок исходной продукции и оборудования, уже в 1970–1980-х гг. стало целесообразным обеспечение глобальной специализации производителей на отдельных видах деятельности. Неслучайно именно с полупроводниковой промышленностью началась глобализация, в результате которой развернутая цепь поставок продукции охватила многие страны мира, в первую очередь страны Дальневосточного региона. Глобализация способствовала удешевлению продукции отрасли и ускорению распространения ИКТ. Считается, что на успешное развитие мировой полупроводниковой промышленности повлияли такие соглашения Всемирной торговой организации (ВТО), как Соглашение об информационных технологиях (Information Technology Agreement – ИТА) от 1996 г., Соглашение по торговым аспектам интеллектуальной собственности (The Agreement on Trade-Related Aspects of Intellectual Property – TRIPS), Соглашение об упрощении процедур торговли (The Trade Facilitation Agreement – TFA) [13].

В результате в отрасли сформировалась сложная и обширная экосистема, состоящая из взаимосвязанных цепочек создания стоимости. Так, в качестве примера в [3] указывается одна американская полупроводниковая компания, которая имеет более 16 тыс. поставщиков по всему миру. Более 7300 ее поставщиков базируются в 46 различных штатах Америки, и более 8500 ее поставщиков находятся за пределами США. Многие из этих поставщиков представляют собой малые предприятия в различных отраслях, которые предоставляют различные товары и услуги, включая химические газы, материалы, строительные услуги, производство пластин услуги, капитальное оборудование, запасные части, системы управления и жизни, вычислительное оборудование, исследования рынка, технические консультации и медиа-услуги.

Отмечается, что Канада, европейские страны и США, как правило, специализируются на разработке полупроводников, а также на производстве продукции

высокого класса. Япония, США и некоторые европейские страны специализируются на поставках оборудования и сырья. Китай, Тайвань, Малайзия и другие азиатские страны, как правило, специализируются на производстве, сборке, тестировании и упаковке. В то же время Канада, Китай, Германия, Индия, Израиль, Сингапур, Южная Корея, Великобритания и США являются крупными центрами исследований и разработок в области полупроводников.

Важным преимуществом участия в глобальной цепочке создания стоимости является положительное влияние, которое оно оказывает на общее экономическое развитие каждой участвующей страны. С другой стороны, в [3] указывается, что в цепочке создания стоимости более 50 точек, в которых на один регион приходится более 65% мирового рынка, что повышает риски нарушения поставок. Так, около 75% мировых мощностей по производству полупроводников, сосредоточено в Китае и Восточной Азии, регионе, который значительно подвержен высокой сейсмической активности и геополитической напряженности. Более того, 100% самых передовых (менее 10 нанометров) производственных мощностей в мире по производству полупроводников в настоящее время расположены на Тайване (92%) и в Южной Корее (8%).

Кроме того, доля компаний США составляет более 50% мирового рынка в 5 основных категориях технологического оборудования для производства полупроводников (инструменты для нанесения покрытий, сухое/влажное травление и очистка, оборудование для легирования, системы управления процессами, тестеры и др.). На Японию приходится более 70% рынка производства фоторезиста, жизненно важного для процесса литографии. Рыночная доля нидерландской компании ASML на мировом рынке установок экстремальной ультрафиолетовой фотолитографии EUV, необходимых для производства современных микросхем с размером менее 7 нанометров, составляет практически 100%.

Неравномерность географического размещения производств повышает риски возникновения нештатных ситуаций и значительных объемов затрат, необходимых на их устранение. В [3] приводятся следующие примеры экстремальных ситуаций, повлиявших на рынок полупроводников.

Так, взрыв на заводе Sumitomo Chemical в Японии в 1993 г. повлиял на 60% мировых поставок эпоксидной смолы. В результате этого спотовые цены на микросхемы памяти DRAM на рынке США выросли в среднем с 30 до примерно 80 долл. за мегабайт (долл/МВ).

Сильное землетрясение в центре Тайваня в сентябре 1999 г. вызвало шестидневную остановку Научного парка Синьчжу из-за отключения электроэнергии. В результате цены на микросхемы памяти выросли втрое. При этом рухнули акции

компаний-производителей электроники, а компании IBM, Hewlett Packard, Intel и Xerox, входившие в то время в список Fortune 100, потеряли от 18 до 40% своей стоимости.

В 2011 г. в Японии произошло сильное землетрясение, за которым последовали цунами и авария на АЭС Фукусима. Катастрофа затронула 25% мирового производства кремниевых пластин и 75% мировых поставок перекиси водорода.

Отмечаются и геополитические риски, связанные с обострением отношений между государствами, чьи компании присутствуют в цепи поставок. Так, в 2019 г. введение Японией контроля на экспорт в Южную Корею фторированного полиимида, фоторезистов и фтористого водорода (на Японию приходится соответственно 90, 70 и 44% мирового производства этих материалов) привело к существенному уменьшению объемов производства и экспорта полупроводников южнокорейскими компаниями [14].

Вместе с тем, в [3] считают, что альтернативный вариант развития отрасли с полностью самодостаточными местными цепочками поставок в каждом регионе потребует не менее 1 трлн долл. дополнительных первоначальных инвестиций и приведет к общему увеличению цен на полупроводники на 35–65%, и, в конечном счете, к более высокой стоимости электронных устройств для потребителей. Угроза появления избыточных мощностей и производств из-за перехода к автаркии еще более усугубит ситуацию. Указывается, что именно из-за глобализации отрасли удалось устранить неблагоприятные для отрасли последствия природных катастроф.

1.3. Роль США и Китая в глобализации полупроводниковой промышленности

США обладают развитой полупроводниковой промышленностью и пока еще определяют научно-технический прогресс и перспективные направления отрасли. В полупроводниковой промышленности занято почти четверть миллиона сотрудников в Соединенных Штатах, а объем продаж американских полупроводниковых компаний в 2020 г. составил 208 млрд долл. [15].

Однако в результате глобализации, предпринятой в конце прошлого столетия по инициативе деловых кругов и администрации этой страны, за последние 30 лет произошло снижение удельного веса США в производстве полупроводников с 37% в 1990 г. до 12% в 2020 г. Вместе с тем, на долю американских компаний, отделения которых разбросаны в разных частях мира, все еще приходится свыше 50% производства интегральных схем (ИС) [5, 16].

После возникшего в 2020 г. дефицита на рынке полупроводников появились предложения по стимулированию производства полупроводниковой продукции в США с целью достижения более диверсифицированного географического размещения производств и снижения риска серьезных перебоев в поставках. Так, в начале 2021 г. Конгресс принял закон, призывающий к федеральным стимулам для увеличения производства микросхем и инвестиций в исследования полупроводников в США для удовлетворения этого растущего спроса [17].

В отчете Комиссии национальной безопасности по искусственному интеллекту (NSCAI) содержался призыв к реализации национальной стратегии в области микроэлектроники, оживлению производства и активизации исследований в области микроэлектроники. Предлагалось осуществить федеральную поддержку исследований и производства полупроводников в объеме 35 млрд долл. (в том числе исследований на сумму 12 млрд долл), а также ввести налоговый кредит для инвестиций в отрасли в размере 40% [18].

Наконец, в июне 2020 г. Конгресс США принял Закон об инновациях и конкуренции (U.S. Innovation and Competition Act – USICA), в котором предусматривается выделение федеральных инвестиций на сумму 52 млрд долл., для проведения исследований и разработки, а также производства в полупроводниковой промышленности страны [19].

Следует также отметить активность, проявляемую Китаем в налаживании собственного производства полупроводников. В настоящее время эта страна является крупнейшим потребителем полупроводников (свыше 30% мирового потребления), но при этом она производит менее 16% от объема потребления этой продукции. В связи с разразившейся в 2018–2019 гг. торговой войной с США, на поставку в китайские компании (включая Huawei) оборудования для фабрик и чипов с малым топологическим размером (14 нм и ниже) были наложены санкции. В результате стране не удалось выполнить задания планов и инициатив, принятых в 2015–2015 гг. Однако в недавно принятом 14-м пятилетнем плане на 2021–2025 гг. намечено обеспечить 70% потребления чипов собственной продукцией. Ввиду того, что многие аналитики считают достижение этой цели маловероятной, в последнее время на Западе усилились опасения относительно обострения ситуации вокруг Тайваня [см. <https://semiengineering.com/china-accelerates-foundry-power-semi-efforts/> и др.].

1.4. Некоторые перспективные направления развития полупроводниковой отрасли

В качестве перспективных направлений использования полупроводниковой продукции консалтинговые компании выделяют автомобильную промышленность и разработки, связанные с искусственным интеллектом. Предполагается также расширение рынка составных (сложных) полупроводников.

1.4.1. Автомобильная промышленность

Отмечается, что в ближайшей перспективе автомобильная электроника и промышленная электроника станут наиболее быстрорастущими сегментами рынка полупроводниковой промышленности

По данным консалтинговой компании Deloitte стоимость полупроводниковых компонентов электронных систем автомобилей (таких, как микроконтроллеры, датчики, память и др.) в расчете на 1 автомобиль выросла с 50 долл в 2000 г. до примерно 475 долл. в 2020 г.. Ожидается, что к 2022 г. эта цифра составит около 600 долл. за автомобиль (табл. 1.2) [21].

Таблица 1.2

Средние затраты на полупроводниковые компоненты
в расчете на 1 автомобиль, долл. [21]

| 2000 | 2010 | 2020 | 2022 |
|------|------|------|------|
| 150 | 300 | 475 | 600 |

1.4.2. Искусственный интеллект (ИИ)

В ближайшие 10 лет одним из главных драйверов роста полупроводниковой промышленности, по мнению консалтинговых организаций, может стать искусственный интеллект (ИИ). В настоящее время ИТ-компании используют в основном 2 типа микросхем при работе с нейронсетевыми программами искусственного интеллекта (ИИ), основанными на использовании больших данных. Так, для ускорения обработки данных и снижения энергоемкости в Google используются микросхемы ASIC, в то время. Как в Microsoft, Xilinx, Baidu и Amazon – микросхемы FPGA.

Ожидается, что рынок полупроводников, связанных с искусственным интеллектом, вырастет с 6 млрд долл в 2019 г. до более чем 30 млрд долл. в 2022 г. со среднегодовым темпом роста почти 50% [2]. При этом в 2022 г. рынок микросхем для ИИ составит более 12% от общего рынка ИИ [21].

1.4.3. Рынок составных (сложных) полупроводников

Составные полупроводники, являющиеся прямозонными (англ. indirect bandgap, в которых переход электрона между зоной проводимости и валентной зоной не сопровождается изменением импульса (так называемый, прямой переход) используются в обычной микроэлектронике, аналоговых интегральных схемах и цифровых интегральных схемах. Они устойчивы к высоким температурам, и характеризуются большим быстродействием и наличием оптоэлектронных свойств. В связи с перечисленными особенностями составных полупроводников, они имеют широкую сферу применения. Так, в автомобильной промышленности составные полупроводники используются для силовых устройств в таких сегментах, как зарядная инфраструктура, а также аккумуляторные и гибридные электромобили. В светотехнике составные полупроводники используются для изготовления светодиодов (LED).

Спрос на составные полупроводники в перспективе, по оценкам консалтинговых компаний, будет увеличиваться. Предполагается, что на развитие рынка составных полупроводников будет, в частности, оказывать значительное влияние прогресс во внедрении технологий беспроводной мобильной связи 5G и следующих поколений, характеризующихся высокой скоростью и малым временем задержки. Отмечается также, что важный фактор роста спроса на составные полупроводники в телекоммуникационной отрасли связан с появлением смартфонов на солнечных элементах на основе GaAs, а также общим увеличением количества смартфонов [22].

Будет продолжаться использование составных полупроводников (например, InP и GaAs) в таких областях, как авионика и радиолокационные системы, а также в военных и аэрокосмических приложениях. Сейчас переключатели на основе GaAs используются, например, в военных тепловизионных системах, работающих в средневолновом (3–5 мкм) диапазоне инфракрасной (ИК) области оптического спектра (MWIR-Middle Wave Infrared) [23].

В целом, ожидается, что мировой рынок составных полупроводников вырастет с 32,0 млрд долл. в 2020 г. до 43,4 млрд долл. к 2025 г. при среднегодовом темпе роста 6,3%. Одним из основных факторов роста этого рынка является растущий спрос и внедрение GaN и SiC в полупроводниковой промышленности [24].

1.5. Развитие рынка полупроводников в перспективе

В связи с неопределенной обстановкой, вызванной пандемией и усилением международной напряженности, существуют объективные трудности при разработке сценариев развития рынка полупроводников даже на краткосрочную перспективу. В настоящей работе на основе простой экономико-статистической модели

построены возможные сценарии развития глобального рынка полупроводников до 2030 г.

В модели исследуются такие показатели, как объем выпуска пластин (wafers) $W_{\text{сл}}$ ИС (приведенных к диаметру 200 мм) в млн ед., объем производимой продукции (доходы или revenues) всего Q_{tot} , в том числе как в дизайнерских компаниях (fables) Q_{FL} , так и в компаниях -производителях пластин с микросхемами (foundries) Q_{F} , в млрд долл. В качестве базовых значений для 2022 г. взяты оценки консалтинговой компании IC Incight. Во всех сценариях в качестве экзогенной переменной принят темп прироста объема выпуска пластин, который определяет динамику остальных показателей рынка.

В связи с тем, что имеет место достаточно надежная линейная зависимость между объемами производства пластин и доходами от производства ИС всего ($Q_{\text{tot}} = 1,5 W$; $R^2 = 0,995$), в трех сценариях использовались одинаковые для этих показателей темпы прироста показателей W и Q_{tot} , равные $\lambda = 0,05$; $\lambda = 0,065$ и $\lambda = 0,08$ (эти значения близки к фактическим величинам темпов прироста исследуемых показателей за разные периоды времени в 2000–2020 гг.).

Связь между объемами производимой продукции (доходами) всего Q_{tot} , и в дизайнерских компаниях (fables) Q_{FL} описывается соотношением:

$$Q_{\text{FL}t} = 0,643 Q_{\text{FL}t-1} + 0,117 Q_{\text{tot}t}, \quad R^2 = 0,996, \quad (1.1)$$

(1,189) (0,048)

а между объемами производимой продукции в дизайнерских компаниях (fables) Q_{FL} , и в компаниях – производителях пластин с микросхемами $Q_{\text{F}t}$ – простым линейным соотношением:

$$Q_{\text{F}t} = 0,516 Q_{\text{FL}t}, \quad R^2 = 0,996 \quad (1.2)$$

(0,011)

(в скобках указаны значения среднеквадратичных ошибок оценок).

Оценки объемов рынка ИС на среднесрочную перспективу (в ценах 2020 г.), полученные на основе соотношений (1.1)–(1.2) для 3-х сценариев, приведены в табл. 1.3–1.5.

Таблица 1.3

Развитие рынка ИС при первом сценарии, млрд долл.
(темп прироста $\lambda = 0,05$)

| | Доходы всего | Доходы фирм-дизайнеров | Доходы фирм-производителей пластин с микросхемами |
|------|--------------|------------------------|---|
| 2020 | 404,4 | 132,6 | 68,4 |
| 2025 | 654,2 | 194,9 | 100,6 |
| 2030 | 834,9 | 252,3 | 130,2 |

Таблица 1.4

Развитие рынка ИС при втором сценарии, млрд долл.
(темпы прироста $\lambda = 0,065$)

| | Доходы всего | Доходы фирм-дизайнеров | Доходы фирм-производителей пластин с микросхемами |
|------|--------------|------------------------|---|
| 2020 | 404,4 | 132,6 | 68,4 |
| 2025 | 682,6 | 200,0 | 103,2 |
| 2030 | 935,2 | 276,6 | 142,7 |

Таблица 1.5

Развитие рынка ИС при третьем сценарии, млрд долл.
(темпы прироста $\lambda = 0,08$)

| | Доходы всего | Доходы фирм-дизайнеров | Доходы фирм-производителей пластин с микросхемами |
|------|--------------|------------------------|---|
| 2020 | 404,4 | 132,6 | 68,4 |
| 2025 | 711,9 | 205,2 | 105,9 |
| 2030 | 1046,0 | 303,1 | 156,4 |

1.5.1. Роль Тайваня в надежности обеспечения поставок

К настоящему времени Тайвань превратился в регион с высокоразвитой высокотехнологичной промышленностью. В 2021 г. суммарный годовой доход 800 публичных компаний Тайваня превысил 900 млрд долл. Особая роль на острове принадлежит отрасли ИКТ. Так, суммарный доход только 7 крупнейших тайваньских компаний ИКТ лишь за III квартала 2021 г. составил более 312 млрд долл. [25, 26].

Бурному экономическому развитию острова способствовало вступление Тайваня, наряду с Китаем и Южной Кореей, в ВТО в 2001 г. Местные аналитики отмечают, что прошедшие 20 лет были для острова «золотыми днями», которые однако закончились в связи с обострением проблем в мировой экономике.

Особое место в развитии высокотехнологичного производства Тайваня занимает основанная в 1987 г. компания TSMC (Taiwan Semiconductor Manufacturing Company Limited). Компания, в которой $\frac{3}{4}$ акций принадлежит иностранным акционерам и в которой трудится 56 тыс. сотрудников, является доминирующим участником рынка производства пластин с микросхемами (Pure-Play Foundries). В 2020–2021 гг. ее удельный вес в производстве пластин с микросхемами превышал 20%, а рыночная доля в доходах всего фабричного сектора (включая IDM-компании) превышала 55%. Большую часть потребителей полупроводниковой продукции составляют дизайнерские фирмы США (более 60%, в том числе Apple – более 20%) и Китая (около 29%). Такое распределение потребителей во многом объясняется высоким

инновационным уровнем выпускаемой компанией продукции. Так в 2020 г. доля доходов компании от продажи пластин с микросхемами, имеющими топологический размер 16 нм и менее, составляла 58%. В настоящее время идет подготовка к запуску процессов с размерами 3 и 2 нм. О технологических успехах TSMC свидетельствует тот факт, что компания Intel, лидер в производстве универсальных процессоров, столкнувшаяся с трудностями освоения техпроцесса 10 нм, предполагает разместить заказы на тайваньских фабриках [27].

Достижения TSMC обусловлены значительным вниманием, уделяемым с начала 2000-х гг. в компании исследованиям и разработкам, на которые расходуется более 8% доходов. В компании были выделены 2 параллельные группы исследователей, одна из которых в течение 18–24 месяцев работала над разработкой и освоением технологий с размерами схем 28 нм, затем 16 нм и далее 7 и 3 нм, а другая – одновременно над разработкой и освоением технологий с размерами схем 20 нм, затем 10 нм и далее 5 нм. Исследователи компания проводили большую работу по освоению и доработке установок экстремальной ультрафиолетовой фотолитографии EUV, необходимых для производства современных микросхем с размером менее 7 нм (в настоящее время на TSMC поставлена половина этих установок, производимых компанией ASML, на поставку которых для других стран распространяются санкции) [28].

В результате, за счет высокой эффективности исследований и разработок, и, как следствие, значительного увеличения наукоемкости производства в расчете на одну пластину, компания добилась существенных конкурентных преимуществ над другими участниками рынка.

Вышеприведенные факты подчеркивают значимость компании как для инновационного развития отрасли, так и для обеспечения надежности цепи поставок. Последнее связано с тем, что объем доходов дизайнерских компаний, зависящих от поставок продукции TSMC, примерно в 2 раза превосходит объем доходов тайваньской компании. Прекращение работы этой компании в силу каких-либо экстремальных причин будет иметь катастрофические последствия для всей отрасли. Так, расчеты на основе вышеприведенной модели (1.1)–(1.3) показывают, что при прекращении работы TSMC в 2024 г. (по нашим оценкам доходы компании могут составить не менее 70 млрд долл.) доходы производителей ИС в целом уменьшатся на 140 млрд долл. Для восстановления к 2028 г. фабрик и прерванного выпуска продукции компании потребуется, по самым скромным оценкам, израсходовать в 2024–2027 гг. капитальных вложений на сумму свыше 133 млрд долл. При таком отклонении, например, от сценария 1, потери за 2024–2030 гг. фирм-дизайнеров могут составить около 600 млрд долл., а всей отрасли – более 910 млрд долл. Об отклонении динамики

доходов в этом случае (т.е. при выходе TSMC из эксплуатации) от первого сценария дает представление рис. 1.2. Значительное снижение объемов производства полупроводниковой продукции в краткосрочной перспективе (до 2025 г.) повлечет спад производства в смежных отраслях и усилит инфляцию, что будет иметь катастрофические последствия для мировой экономики.



Рис. 1.2. Динамика объемов мирового рынка ИС в сценарии 1 и при отклонении от него (выход TSMC) в млрд долл.

1.6. Моделирование динамики показателей рынка высокотехнологичной продукции с догоняющим участником (на примере фабричного производства чипов)

В связи с усилением международной напряженности, распространившейся в конце 2019 – начале 2020 гг. по всему миру пандемией и глобальным экономическим кризисом, появились тенденции перехода мировой экономики к деглобализации. Существенно возросли риски несоблюдения основополагающих принципов торговли, закрепленных в постановлениях и соглашениях Всемирной торговой организации (ВТО). Во многом этому способствуют торговые и технологические войны между ведущими экономическими странами и компаниями (достаточно указать на торговую войну, разразившуюся в 2018 г. между США и Китаем). Особо остро эти явления проявляются на рынке полупроводников, определяющем в настоящее время технологический прогресс всей мировой экономики. Так, уже с 2018 г. США стали накладывать разного рода санкции на поставку в Китай современного оборудования для производства чипов (с топологическим размером менее 14 нм).

Вышеуказанные процессы требуют переосмысления и развития традиционных методов микроэкономики, особенно теории организации промышленности. В частности, в условиях сохраняющейся взаимозависимости экономических агентов друг от друга, усиливается необходимость учета действий некоторых участников рынков, не согласующихся с базовыми принципами рыночной экономики, а также реакции на эти действия со стороны других участников рынка, формально приверженных твердому соблюдению этих принципов, но также не соблюдающими их. Особую актуальность приобретает исследование взаимодействий таких экономических агентов на олигополистических рынках высокотехнологичной продукции, определяющих перспективы развития человеческого общества. Одним из таких рынков является рынок производителей пластин с микросхемами (Pure Pay Foundries, см. п. 1.2.1.)

В настоящем п. рассматривается динамика структуры таких рынков в условиях вхождения на них новых, «догоняющих» участников, руководствующихся неэкономическими критериями.

1.6.1. Модель динамики показателей олигополистических рынков

Проводимый в настоящей главе анализ основан на использовании агрегированной динамической модели рационального поведения участников олигополистического рынка производителей пластин с микросхемами, в виде линейной динамической игры по Нэшу–Курно с квадратичным критерием. Участниками рынка являются N фирм-олигополистов и $N+1$ -й догоняющий участник [29, 30]. Центральным блоком модели является следующая динамическая зависимость, связывающая объемы товарного производства фирм Q_{it} со входной переменной u_{it} (инвестициями) с помощью передаточных функций $W_i(z)$, i – индекс фирмы, $i=1, 2, \dots, N$, t – индекс года:

$$Q_{it} = W_i(z) u_{it} + Q_{0it}, \quad (1.3)$$

где $W_i(z) = \frac{\gamma_{0i}z + \gamma_{1i}}{(z - \lambda_i)^2}$, $i=1, 2$; $t=0, 1, \dots$; z – оператор сдвига: $zx_t = x_{t+1}$;

$\gamma_{0i}, \gamma_{1i}, \lambda_i$ – параметры, Q_{0it} – слагаемое, характеризующее начальные условия.

Другой блок модели – обратная функция спроса, представляющая собой линейную зависимость цен олигополистов на рынке P_{it} от уровней наукоемкости их

продукции $rdw_{it} = \frac{R \& D_{it}}{Q_{it}}$ ($R \& D_{it}$ – объем затрат на исследования и разработки в i -

й компании) и от суммарного объема предложения продукции $Q_t = \sum_{i=1}^{N+1} Q_{it}$:

$$P_{it} = a_0 + a_i \frac{R \& D_{it}}{Q_{it}} - bQ_t, \quad (1.4)$$

где $a_i, b, i = 0, 1, 2, \dots, N$ – параметры.

Предполагается, что объем производства продукции догоняющим участником Q_{N+1t} связан с суммарным объемом производства в олигополии соотношением:

$$Q_{N+1t} = \lambda_{N+1} Q_{N+1,t-1} + g(1 - \lambda_{N+1}) Q_{t-1}, \quad (1.5)$$

где $0 < \lambda_{N+1} < 1$. Цель догоняющего участника рынка, не испытывающего ограничений в финансово-экономических, материальных и трудовых ресурсах, состоит лишь в обеспечении доли на рынке, равной $MS_{N+1} = g$.

Олигополисты в каждый момент времени τ максимизируют чистую текущую стоимость (NPV) или прибыль:

$$J_{it} = \sum_{t=\tau}^{\tau+T_{pi}} \beta_i^t \left[\chi_i P_{it} Q_{it} - q_i u_{it} - \frac{1}{2} \rho_i u_{it}^2 \right] \rightarrow \max_{u_{it}}, \quad (1.6)$$

где $\beta_i = 1 / (1 + r_i)$ – дисконтирующий множитель, соответствующий дисконт-фактору r_i ; P_{it} – цена единицы продукции компании; χ_i – доля прибыли в доходах компании; $q_i = 1$ или $q_i = 0$ – в зависимости от используемого i -м олигополистом критерия; $\frac{1}{2} \rho_i u_{it}^2$ – затраты регулирования, характеризующие инвестиционные возможности олигополиста; $\rho_i > 0, i = 1, 2, \dots, N$; T_{pi} – период скользящего планирования (для упрощения записи формул ставки налогов приняты равными нулю); $\tau = 1, 2, \dots, T, T$ – период планирования. Управляющими переменными для олигополистов в модели являются объемы производственных инвестиций $-u_{it}, i = 1, 2, \dots, N$.

В данном п. расчеты оптимальных по Нэшу-Курно разомкнутых (open-loop) стратегий олигополистов проведены с использованием обобщенных (generalized) матричных уравнений Риккати (подробнее см., например, [31], а также главу 2).

1.6.2. Результаты расчетов

В настоящем п. рассматривается модель олигополии, состоящей из 4-х ведущих фирм рынка полупроводниковых пластин (1. TSMC; 2. UMC; 3. SMIC; 4. Global Foundries), на которые приходится примерно 80% доходов чисто фабричных

компаний (Pure Pay Foundries), а также одного догоняющего участника. Параметры зависимостей (1.3)–(1.4) для исследуемой олигополии, полученные методом наименьших квадратов (соответственно нелинейным и линейным) представлены в табл. 1.6–1.7.

Таблица 1.6

Параметры передаточных функций компаний квадрополии

| № компании | λ_i | γ_{0i} | γ_{1i} |
|------------|-------------|---------------|---------------|
| 1 | 0,434 | 0,743 | 0 |
| 2 | 0,755 | 0,125 | 0 |
| 3 | 0,643 | 0 | 0,533 |
| 4 | 0,678 | 0,3917 | 0 |

Таблица 1.7

Параметры обратной функции спроса (1.4) и входящие в критерии (1.6)

| № компании | a_0 | a_{1i} | b | χ_i | q_i |
|------------|-------|----------|--------|----------|-------|
| 1 | 0,890 | 5,620 | 0,0029 | 0,479 | 1 |
| 2 | 0,890 | 2,644 | 0,0029 | 1,000 | 0 |
| 3 | 0,890 | 0 | 0,0029 | 0,196 | 0 |
| 4 | 0,890 | 0 | 0,0029 | 0,227 | 0 |

В расчетах приняты постоянные значения наукоёмкости продукции $rdw_i = \frac{R \& D_{it}}{Q_{it}}$, составляющие для компаний в базовом варианте соответственно 0,1324, 0,0992, 0,0516 и 0,1189 единиц. Принято также, что цель догоняющего участника состоит в достижении 30%-й рыночной доли ($g = 0,3$), причем в базовом варианте параметр λ в соотношении (1.5) составляет $\lambda = 0,5$.

Ниже представлены результаты расчетов влияния на структуру рынка таких параметров модели, как скорость освоения производства догоняющим участником рынка (на рисунках – follower), наукоёмкость и эффективность производства продукции у фирмы-лидера (TSMC, на рисунках – leader), а также некоторых подходов к противодействию догоняющего участника фирмой-лидером.

1.6.2.1. Анализ влияния скорости освоения производства догоняющим участником

Скорость освоения производства догоняющим участником рынка в данной модели с соотношением (1.5), описывающим динамику его производства, характеризуется значением параметра λ (при этом средний лаг отставания составляет

$T_L = \frac{1-\lambda}{\lambda}$ лет). Очевидно, чем выше значение λ , тем сильнее отставание

догоняющего участника от остальных фирм, или, что равносильно, тем ниже скорость освоения им производства. При этом вытеснение остальных олигополистов происходит тем интенсивнее, чем меньше значение параметра λ . Этот факт иллюстрируют рис. 1.3–1.5 (ниже на всех рисунках $L = \lambda$). Так, при $\lambda = 0,75$ максимальный объем производства компании-лидера заметно больше, чем при $\lambda = 0,25$. При этом значении λ имеет место существенное отставание догоняющего участника рынка от достижения цели (рис. 1.3–1.5).

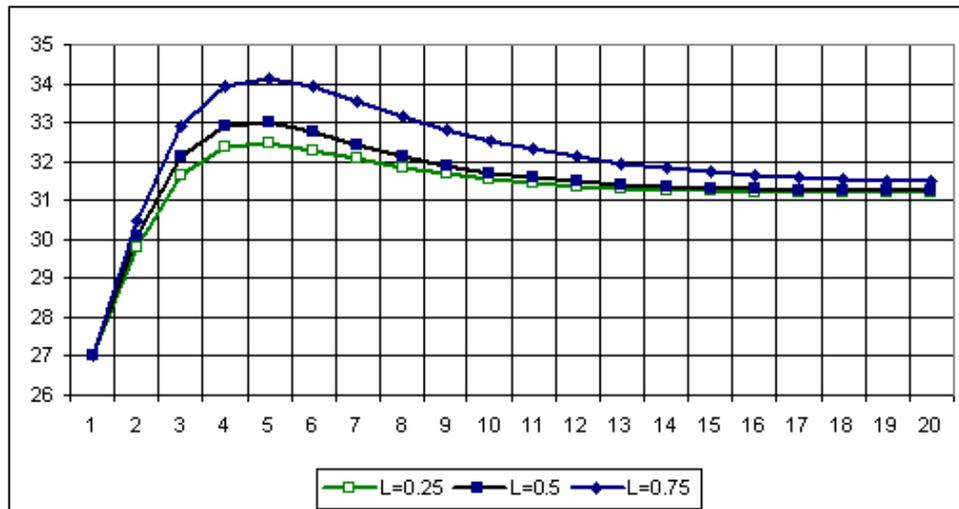


Рис. 1.3. Динамика объемов производства компании-лидера при разных лагах отставания догоняющего участника рынка от цели

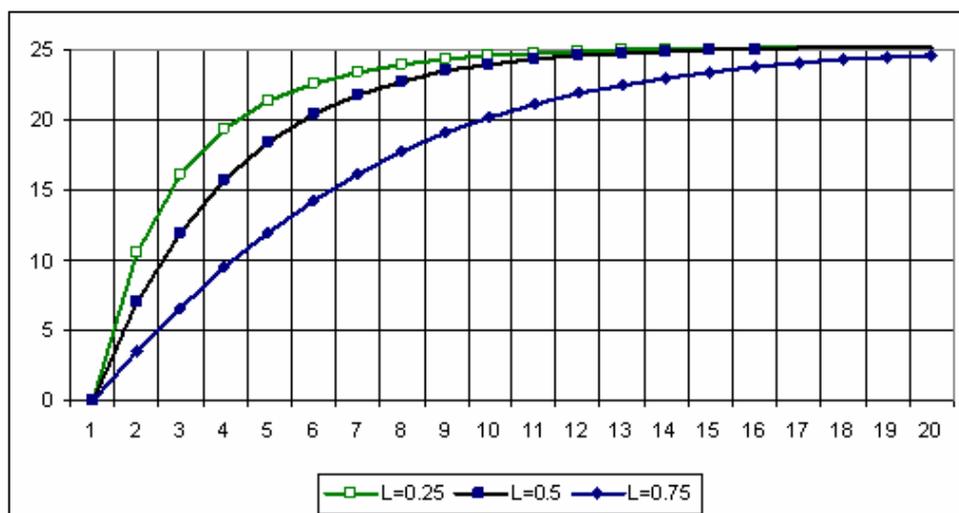


Рис. 1.4. Динамика объемов производства догоняющего участника рынка при разных лагах его отставания от цели

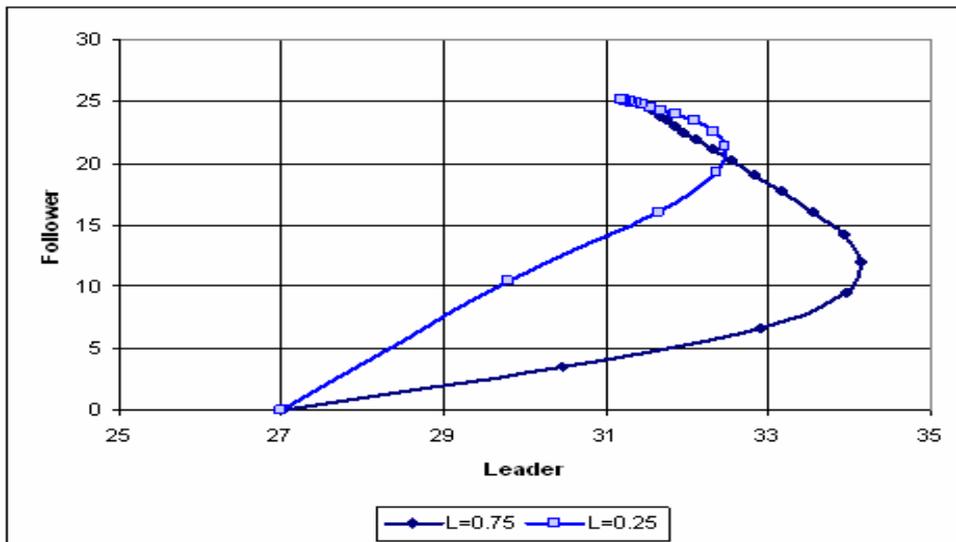


Рис. 1.5. Фазовый портрет, характеризующий связь объемов производства лидера и догоняющего участника рынка при разных лагах отставания последнего от цели

Следует отметить также, что вхождение на рынок догоняющего участника сопровождается существенным уменьшением объемов производства и рыночной доли старых фирм-олигополистов (рис. 1.6–1.7) по сравнению со случаем отсутствия нового, $N+1$ -го участника. За 20-летний период это уменьшение в базовом варианте составляет более 11%.

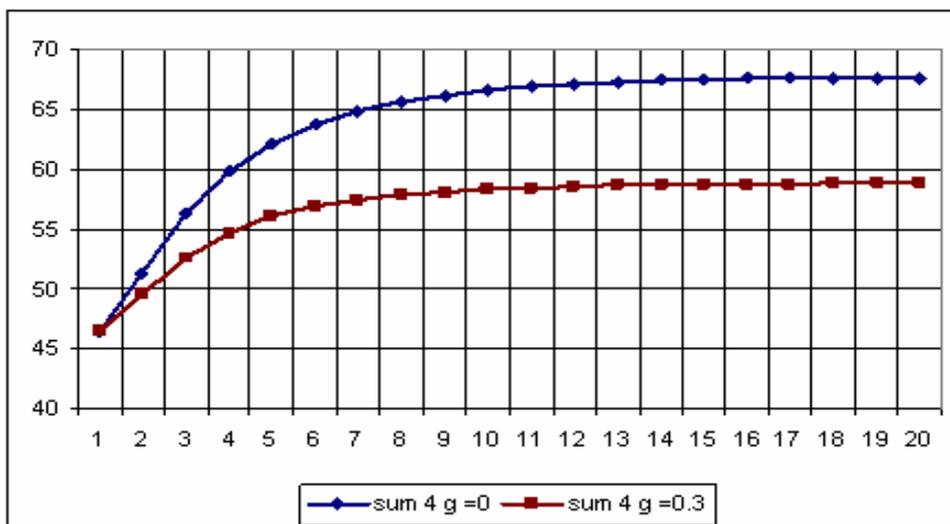


Рис. 1.6. Динамика объемов производства 4-х олигополистов при отсутствии ($g = 0$) и наличии ($\lambda = 0,5; g = 0,3$) догоняющего участника рынка

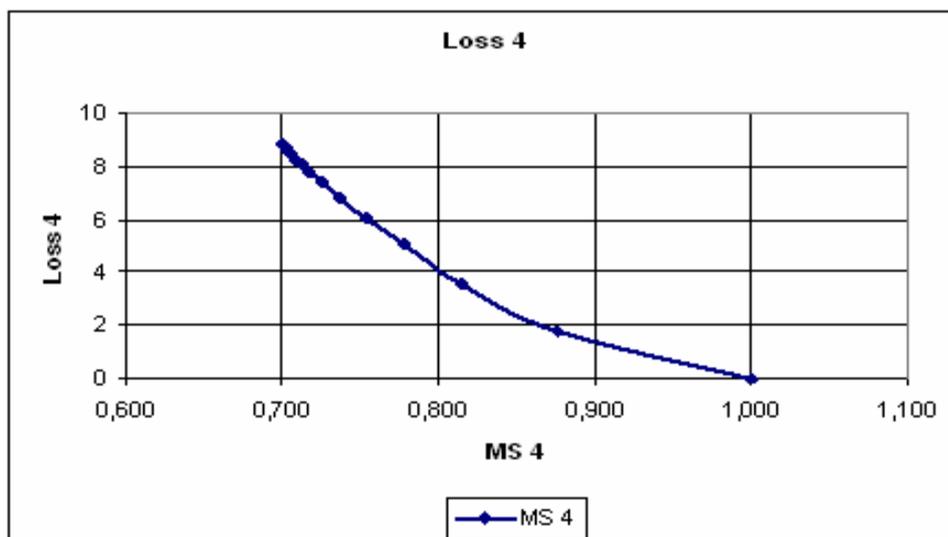


Рис. 1.7. Связь объемов теряемой продукции (Loss) с уменьшением рыночной доли 4-х олигополистов (MS 4) при вхождении на рынки догоняющего участника ($\lambda = 0,5$; $g = 0,3$)

1.6.2.2. Анализ влияния показателей эффективности лидера на динамику его производства

Одним из факторов уменьшения потерь олигополистов, связанных с вхождением на рынок нового участника, является повышение эффективности их производства. В рамках используемой модели показателями эффективности являются уровень наукоемкости продукции rdw_i и доля прибыли в доходах компаний χ_i . Часто, но, конечно, не всегда, уровень наукоемкости высокотехнологичной продукции отражает ее инновационность и большую востребованность у потребителей, хотя и при большей цене (в настоящее время эта закономерность проявляется в производстве интегральных схем с малым топологическим размером). В то же время увеличение доли прибыли в доходах компаний-олигополистов однозначно ведет к снижению чувствительности их показателей к всевозможным возмущениям и изменениям на рынке.

На рис. 1.8–1.11 представлены результаты расчетов на основе модели (1–4) динамики объемов производства и рыночной доли компании-лидера и догоняющего участника рынка при разных значениях наукоемкости лидера ($rdw_1 = 0,1324$ и $rdw_1 = 0,15$), а также при разных значениях доли прибыли в доходах компании лидера ($\chi_1 = 0,479$ и $\chi_1 = 0,579$). Расчеты показывают, что при большом лаге отставания догоняющего участника рост наукоемкости продукции лидера приводит в среднесрочной перспективе к большему увеличению объемов его производства

относительно базового уровня, чем при малом лаге (соответственно на 20 и 16%,
рис. 1.8–1.9).

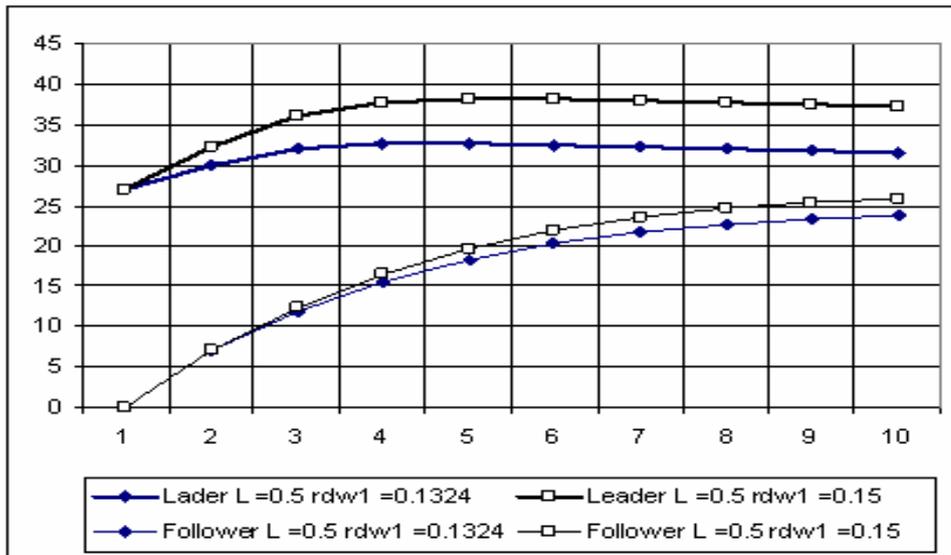


Рис. 1.8. Динамика объемов производства компании-лидера и догоняющего участника рынка при разных уровнях наукоёмкости продукции лидера

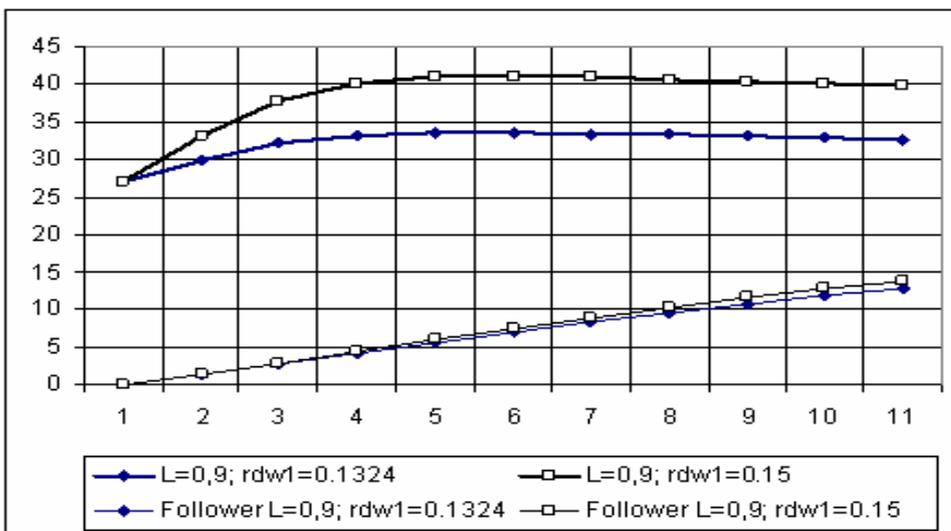


Рис. 1.9. Динамика объемов производства компании-лидера и догоняющего участника рынка при разных значениях наукоёмкости продукции лидера и высоком лаге отставания

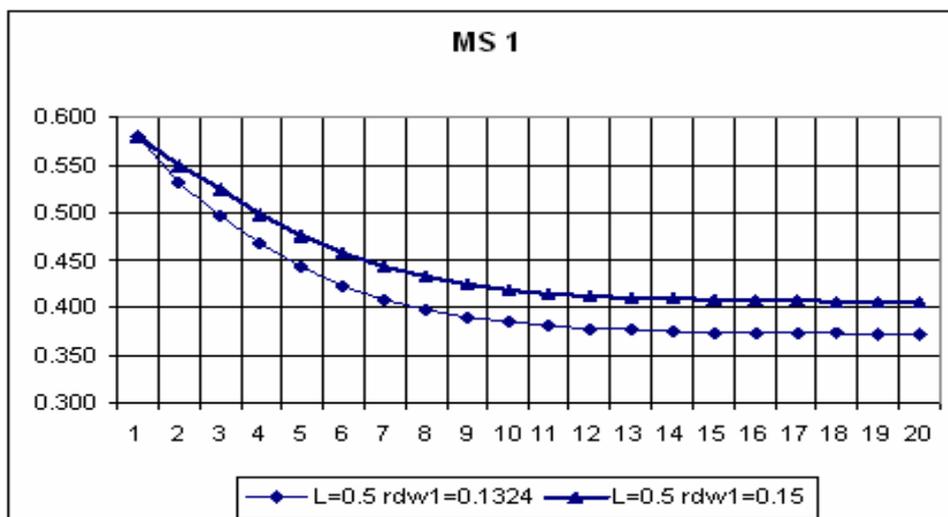


Рис. 1.10. Динамика рыночной доли компании-лидера при разных значениях наукоёмкости продукции лидера

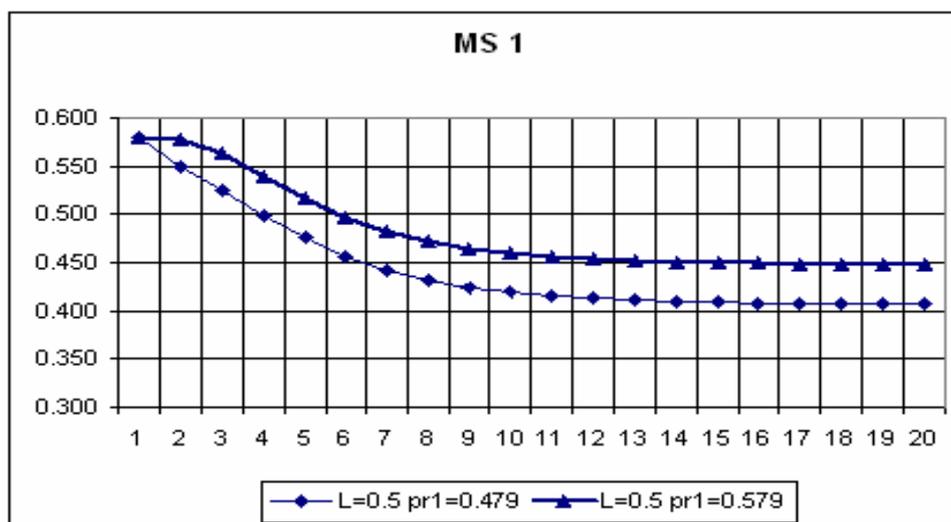


Рис. 1.11. Динамика рыночной доли компании-лидера и догоняющего участника рынка при разных значениях доли прибыли в доходах компании лидера χ_i

1.6.2.3. Анализ эффективности противодействия лидера повышению рыночной доли догоняющего участника

На основе модели (1.3)–(1.6) проводилось также исследование некоторых мер нерыночного противодействия новому участнику со стороны фирмы-лидера (например, закрытие рынка для части его продукции, или недопоставка сырья, материалов, оборудования и др.). Так, вместо (1.5) использовалось соотношение:

$$Q_{N+1t} = \lambda_{N+1} Q_{N+1t} + g(1 - \lambda_{N+1}) Q_t - v_t, \quad (1.5a)$$

где v_t – воздействие на объем предложения продукции нового участника Q_{N+1t} со стороны лидера. В расчетах исследовалось 2 варианта противодействия:

1) пропорциональное объему выведенной на рынок продукции нового участника в момент $t - 1$, т.е. $v_t = f \cdot Q_{N+1,t-1}$;

2) пропорциональное объему продукции фирмы-лидера, т.е. $v_t = f \cdot Q_{it}$, где f – коэффициент.

Проведенные расчеты показывают, что реализация лидером мер по снижению продукции, вводимой новым участником, сопровождается при втором варианте противодействия большим объемом продукции лидера чем при первом варианте (рис. 1.12–1.13).

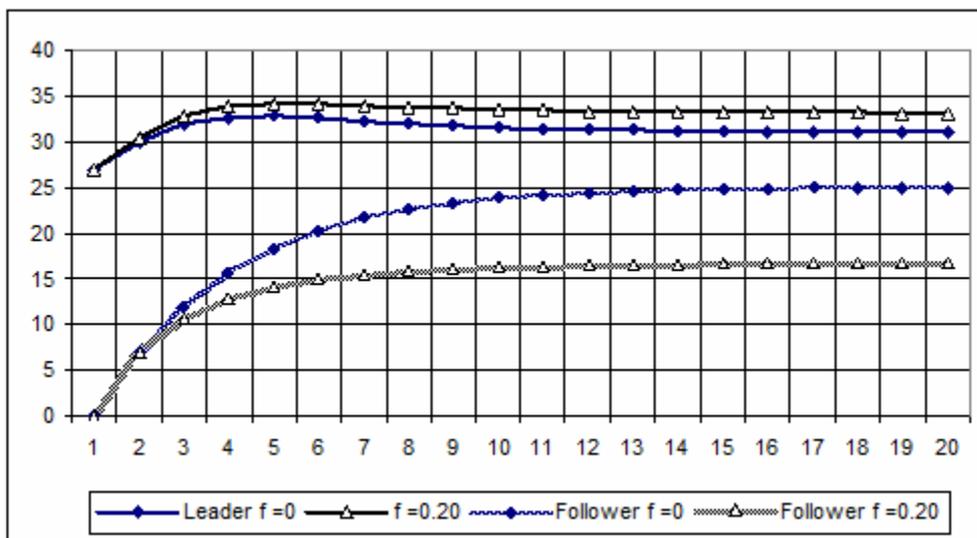


Рис. 1.12. Динамика объемов производства компании-лидера и догоняющего участника рынка при отсутствии ($f = 0$) при первом варианте противодействия ($f = 0,20$) со стороны лидера

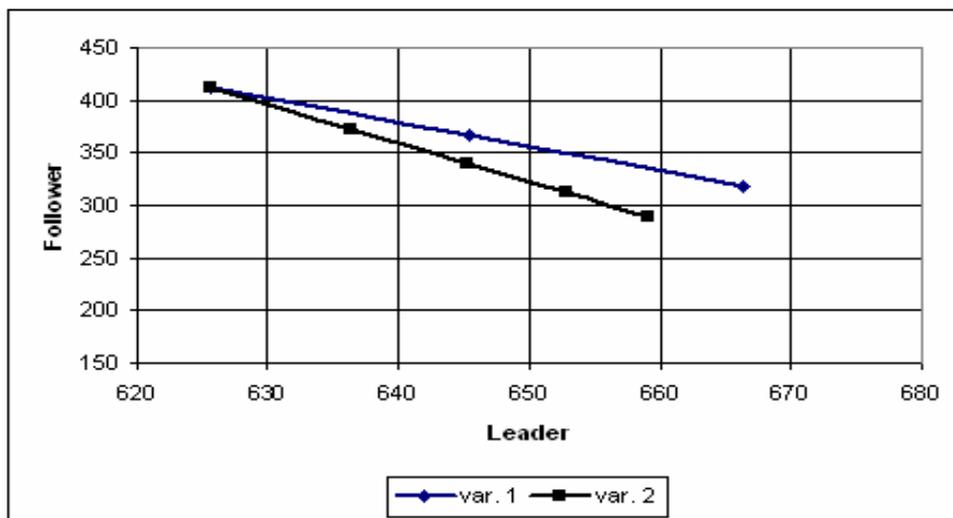


Рис. 1.13. Связь объемов производства компании-лидера и догоняющего участника рынка за 20-летний период при двух вариантах противодействия со стороны лидера (построена при разных значениях коэффициента f)

Полученные результаты расчетов позволяют сделать следующие выводы:

- 1) вытеснение олигополистов новым участником происходит тем интенсивнее, чем меньше лаг отставания последнего от целевых уровней;
- 2) рост наукоемкости продукции лидера приводит в среднесрочной перспективе к большему увеличению объемов его производства относительно базового уровня при большом лаге отставания догоняющего участника, чем при малом лаге;
- 3) более эффективным из рассмотренных вариантов противодействия догоняющему участнику является вариант закрытия рынка для части его продукции в размере, пропорциональном объему продукции фирмы-лидера.

1.7. Заключение

Спрос на полупроводниковую продукцию, обусловлен непрерывным внедрением экономически эффективных инноваций в полупроводниковой отрасли.

В связи с высокой науко- и капиталоемкостью, полупроводниковая отрасль стала одним из первых объектов глобализации.

В связи с усилившейся конкуренцией со стороны дальневосточных производителей и повышением риска нарушения цепи поставок в сложившейся экосистеме, полупроводниковой продукции, США столкнулись с необходимостью локализации полупроводниковых производств на своей территории.

Перспективными направлениями развития полупроводниковой отрасли, по мнению консалтинговых организаций, являются автомобильная промышленность и искусственный интеллект. Вместе с тем, предполагается расширение рынка составных (сложных) полупроводников.

В связи с обострением международной обстановки и пандемией стали отчетливо проявляться негативные явления, связанные с глобализацией, что заставляет развитые страны переходить к локализации полупроводниковых производств на своей территории.

В кратко- и среднесрочной перспективе серьезное нарушение сложившихся структуры отрасли и цепи поставок может привести к катастрофическим последствиям для экономики и к провалу планов цифровизации общества

Рассмотренный подход к моделированию динамики показателей рынка высокотехнологичной продукции может быть использован при прогнозировании реальных рынков микроэлектронной продукции с догоняющими участниками.

Список литературы к главе 1

1. Варшавский Л.Е. «Современные тенденции развития полупроводниковой промышленности» (Концепции. 2021. № 1 (40). С. 44–50).
2. Варшавский Л.Е. «Исследование эволюции рынка высокотехнологичной продукции с догоняющим участником» (Теория и практика институциональных преобразований в России: сб. статей / Ерзнкян Б.Г. (ред.). Вып. 53. М.: ЦЭМИ РАН, 2022. С. 61–70).
3. Strengthening the Global Semiconductor Supply Chain in an Uncertain Era SIA/BCG Report. April 2021. URL: <https://www.bcg.com/publications/2021/strengthening-the-global-semiconductor-supply-chain> (дата обращения 21.05.2021).
4. URL: <https://www.electronicweekly.com/news/business/auto-strongest-semi-market-segment-2020-06/> (дата обращения 25.05.2021).
5. URL: <https://www.electronicweekly.com/news/business/us-holds-50-ic-market-share-2019-06> (дата обращения 29.05.2021).
6. URL: <https://www.eenewseurope.com/news/semiconductor-market-rise-108-percent-2022> (дата обращения 15.02.2021).
7. URL: <https://www.vedomosti.ru/business/articles/2022/01/13/904637-prodazhi-personalnih-kompyuterov> (дата обращения 25.02.2022).
8. URL: <https://www.semiconductors.org/chipmakers-are-ramping-up-production-to-address-semiconductor-shortage-heres-why-that-takes-time/> (дата обращения 25.03.2021).
9. Perry T.S. Morris Chang: Foundry Father. URL: <http://spectrum.ieee.org/at-work/tech-careers/morris-chang-foundry-father>
10. Варшавский Л. Е. Моделирование динамики ключевых показателей рынков компонентов высокопроизводительных вычислительных систем // Труды Института системного анализа Российской академии наук (ИСА РАН). 2017. Т. 67. Вып. 1. С. 12–27.
11. Samsung and TSMC Seeking to Spend Their Way to Worldwide Domination of Advanced IC Technology Research Bulletin March 16, 2021 2021 by IC INSIGHTS, INC. URL: <https://www.icinsights.com/news/bulletins/Samsung-And-TSMC-Seeking-To-Spend-Their-Way-To-Worldwide-Domination-Of-Advanced-IC-Technology/> (дата обращения 29.05.2021).
12. URL: <https://www.notebookcheck.net/TSMC-wants-to-solve-the-global-chip-shortages-with-100-billion-investment-in-new-worldwide-production-facilities.530573.0.html/> (дата обращения 21.05.2021).
13. URL: <https://www.semiconductors.org/trade-in-transition-the-importance-of-re-establishing-u-s-trade-leadership/> (дата обращения 25.05.2021).
14. URL: <https://www.imemo.ru/publications/relevant-comments/text/yapono-yuzhnokoreyskie-raznoglasiya-i-reaktsiya-ssha> (дата обращения 25.05.2021).
15. URL: <https://www.semiconductors.org/global-semiconductor-sales-increase-6-5-to-439-billion-in-2020/> (дата обращения 15.05.2021).
16. URL: <https://www.semiconductors.org/semiconductor-industry-leaders-urge-president-biden-to-prioritize-funding-for-semiconductor-manufacturing-research/> (дата обращения 25.05.2021).
17. URL: <https://www.semiconductors.org/semiconductor-shortage-highlights-need-to-strengthen-u-s-chip-manufacturing-research/> (дата обращения 25.05.2021).
18. URL: <https://www.semiconductors.org/ai-commission-calls-for-federal-investments-in-semiconductor-manufacturing-and-research/> (дата обращения 17.05.2021).
19. URL: <https://www.semiconductors.org/chips/> (дата обращения 19.06.2021).
20. URL: <https://semiengineering.com/china-accelerates-foundry-power-semi-efforts/> (дата обращения 19.02.2022).

21. Semiconductors – the Next Wave | The ever-evolving semiconductor scene. Deloitte April 2019. URL: <https://www2.deloitte.com/cn/en/pages/about-deloitte/articles/pr-semiconductors-the-next-wave-2019.html> (дата обращения 25.05.2021).
21. URL: <https://www.researchandmarkets.com/reports/4986735/compound-semiconductor-market-forecasts-from#relb0-4858405> (дата обращения 22.05.2021).
23. Тарасов В.В., Якушенков Ю.Г. Современное состояние и перспективы развития зарубежных тепловизионных систем // Научно-технический вестник информационных технологий, механики и оптики. 2013. № 3 (85). С. 1–13.
24. URL: <https://www.marketsandmarkets.com/Market-Reports/compound-semiconductor-market-178858112.html> (дата обращения 23.05.2021).
25. URL: <https://www.digitimes.com/news/a20220124VL208.html?mod=2> (дата обращения 1702.2022).
26. URL: <https://www.digitimes.com/news/a20211103VL206.html?mod=2> (дата обращения 1702.2022).
27. URL: <https://www.digitimes.com/news/a20210817VL203.html?mod=2> (дата обращения 1702.2022).
28. URL: <https://www.digitimes.com/news/a20210712VL201.html> (дата обращения 1702.2022).
29. Варшавский Л.Е. Использование методов теории управления для формирования рыночных структур // Компьютерные исследования и моделирование. 2014. Т. 6. № 5. С. 839–859.
30. Варшавский Л.Е. Исследование динамики структуры олигополистических рынков при нерыночных противодействиях сторон // Компьютерные исследования и моделирование. 2021. Т. 13. № 1. С. 219–233.
31. Kwon W.H., Han S.H. Receding horizon control: model predictive control for state models. London, Springer, 2005.

ГЛАВА 2.

МОДЕЛИРОВАНИЕ ДИНАМИКИ КЛЮЧЕВЫХ ПОКАЗАТЕЛЕЙ РЫНКА МИКРОПРОЦЕССОРОВ

2.1. Введение

Развитие высокопроизводительных вычислительных систем (НРС-систем), и особенно суперкомпьютерных технологий, является одним из решающих факторов повышения конкурентоспособности и укрепления национальной безопасности страны. В связи с этим актуальными являются исследования в области анализа и прогнозирования развития рынка НРС-систем и компонентов для них.

К настоящему времени опубликовано относительно небольшое число работ, относящихся к экономическому исследованию и моделированию различных аспектов развития рынков отдельных компонентов для компьютеров (в основном систем памяти и микропроцессоров) и НРС. Так, в работе [1] проведено исследование технологических изменений рынка компонентов для персональных компьютеров. Разработана эконометрическая модель динамики показателей дуополистического рынка микропроцессоров x86, на основе которой построены сценарии поведения ключевых показателей рынка (средней цены, инвестиций, мощностей, рыночной структуры) на 2002–2006 гг.

В работе [2] построена модель, восходящая к пионерской работе Solow (типа *vintage-capital*). Значительное внимание в работе уделяется теоретическому анализу динамики цены на микропроцессоры и взаимосвязи темпов снижения цен на вводимые поколения микропроцессоров с темпами уплотнения чипов в соответствии с, так называемым, «законом Мура».

В статье [3] также на основе модели типа *vintage-capital* исследуются факторы, повлиявшие на динамику технологического прогресса в подотрасли производства микропроцессоров в разные периоды времени.

В статье [4] исследование дуополистического рынка микропроцессоров проводится с учетом того, что драйверами инноваций на рынке являются как конкуренция между дуополистами, так и конкуренция между новым и старым оборудованием, вынуждающая потребителей проводить модернизацию и обновление компьютерного парка. Один из полученных в статье выводов состоит в том, что из-за второго драйвера, монополизация рынка компанией Intel в сложившихся условиях заставила бы последнюю проводить более интенсивную инновационную политику.

В работах [5, 6] исследуется влияние технологических инноваций на структуру рынка суперкомпьютеров, участниками которого являются многопродуктовые

фирмы. При ограниченном числе участников рынка показано наличие положительной связи между уровнем конкуренции и инновационной активностью.

Следует, однако, отметить, что вне поля зрения авторов остаются, например, такие появившиеся в последнее время проблемы, как: возможное расширение сегментов рынка микропроцессоров и других приборов в направлении создания энергоэкономичных компонентов для серверов, центров обработки данных (ЦОД, англ. data centers), Интернета вещей (Internet of Things – IoT) и др. приборов и систем. Кроме того, необходимо подчеркнуть, что по имеющейся у нас информации, большая часть работ экономистов в рассматриваемых областях ориентирована лишь на описание сложившихся явлений и процессов с помощью понятий и концепций современной экономической теории и эконометрики, а не на построение инструментария для формирования прогнозов и сценариев развития рынков, связанных с НРС.

Данная проблема актуальна для нашей страны при построении эффективной стратегии импортозамещения, которая должна учитывать опыт развития ведущих компаний мира. Для ее решения в сложившихся условиях особенно важно иметь правильные оценки динамических характеристик развития компаний-производителей компонентов НРС в условиях конкуренции. Эти характеристики могут быть использованы при прогнозировании и планировании показателей деятельности отечественных предприятий, особенно показателей эффективности производственных инвестиций, а также при формировании конкурентной среды в разных направлениях (сегментах) подотрасли.

В связи с определяющим влиянием элементной базы на производительность НРС-систем и обостряющейся проблемой ее импортозамещения, в настоящей главе значительное внимание уделяется разработке экономико-математической модели развития рынка микропроцессоров для персональных компьютеров (PC) в целом и для серверов, в которых отражены наметившиеся структурные сдвиги (насыщение рынка персональных компьютеров, выход на рынок микропроцессоров для серверов новых производителей и др.).

В связи с отсутствием официальной статистики об объемах производства различных компонентов PC и серверов, при разработке модели приходится ориентироваться на неполные данные разных источников, доступ к которым затруднен. Это усложняет обновление рассматриваемых в главе моделей и прогнозных сценариев, разработанных в 2016 г. и представленных в статьях автора [7–9]. Тем не менее используемый в главе подход к моделированию, основанный на теории динамических игр, может быть полезен и при наличии более свежих данных.

Материал настоящей главы основан на статьях автора [7–9].

2.2. Основные тенденции развития элементной базы НРС (на примере микропроцессоров)

Основу элементной базы НРС-систем составляет продукция микроэлектронной промышленности. Как отмечалось в главе 1, зарождение микроэлектроники связывают с появлением интегральных схем (ИС) в конце 1950-х – начале 1960-х гг. Особую роль в появлении отрасли сыграли потребности и заказы аэрокосмической промышленности. Так, в США важную стимулирующую роль в разработке ИС сыграла работа по созданию ракеты Minuteman. В условиях острого противостояния в период холодной войны, оборонное ведомство этой страны предъявляло повышенные требования к техническим характеристикам авионики для ракетной техники (включая массогабаритные, точность, надежность и др.), которые не могли быть удовлетворены при использовании существовавших тогда дискретных схем. Эти требования побудили молодых тогда сотрудников компаний-разработчиков полупроводниковой продукции (в первую очередь Fairchild, Texas Instruments) искать новые решения по повышению, прежде всего, надежности схем, что в итоге привело к созданию ИС и новой подотрасли (отметим, что наши специалисты старались не отставать от американских [10]).

Быстрое развитие микроэлектроники в 1960-х гг. во многом было связано с возникновением в Кремниевой долине (Silicon Valley) целого кластера новых фирм и организацией их поддержки со стороны венчурных фондов. Прогресс в области миниатюризации ИС позволил наладить выпуск широкой гаммы продукции для промышленного и бытового применения. Появились новые разработки для промышленной и бытовой электроники (модули полупроводниковой памяти, калькуляторы, часы и др.), возникли новые сегменты рынка, на которых в 1970–1980-х гг. развернулась жесткая конкурентная борьба между американскими и, прежде всего, японскими компаниями.

Ряд компаний Кремниевой долины стал прародителем многих других (возникло такое понятие, как spin off). В этом отношении уникальным явлением является компания Fairchild. На базе этой компании с численностью в лучшие годы 1400 чел. возникло множество малых компаний (spinoffs) с численностью 12 500 чел. Некоторые из них (особенно Intel, AMD) превратились впоследствии в гигантов микроэлектроники. Так, изобретение в начале 1970-х гг. в Intel микропроцессора позволило этой компании получить право производить микропроцессоры архитектуры x86 для разработанных в компании IBM персональных компьютеров и впоследствии, отказавшись от выпуска модулей памяти, стать ведущим участником нового рынка [11–13].

Одним из важнейших драйверов развития микроэлектроники с момента ее зарождения стало повышение степени интеграции полупроводниковых схем, что обеспечивало увеличение быстродействия, снижение рассеиваемой энергии и повышение качества и надежности работы схем. Еще в середине 1960-х гг. видным специалистом в области микроэлектроники, одним из основателей компании Intel Г. Муром (H. Moore) была выдвинута широко известная ныне гипотеза об увеличении в 2 раза степени интеграции интегральных схем через каждые 1,5–2 года, впоследствии ставшая императивом последующих изменений в микроэлектронике [14]. Значительное влияние на развитие микроэлектронных схем и устройств оказала вышедшая в 1974 г. работа Р. Деннарда (R. Dennard) с сотрудниками, в которой было показано, что в чипах на основе КМОП-технологии (CMOS) плотность рассеиваемой энергии сохраняется постоянной при одинаковом уменьшении в k раз топологического размера, подводимого к транзистору напряжения, а также при увеличении во столько же раз внутренней частоты устройства (в этом состоит, так называемый закон Деннарда). Следование этой закономерности позволяло вплоть до начала 2000-х гг. добиваться роста производительности устройств (микропроцессоров, памяти) главным образом за счет их минитюаризации и увеличения частоты [15].

Так, за период с 1971 по 2000 г. количество транзисторов в микропроцессорах компании Intel увеличилось почти в 19 тыс. раз, а внутренняя частота – в 30 тыс. раз! Однако уже в середине первого десятилетия XXI в., в связи с форсированием разработчиками роста производительности микросхем путем использования решений, способствовавших несбалансированному с уменьшением размеров схем росту тактовой частоты (в первую очередь это относится к создававшимся микропроцессорам), кремниевая полупроводниковая КМОП-технология стала подходить к границам своего развития. К этому времени стали остро ощущаться проблемы, связанные с повышением энергоемкости электронных приборов при уменьшении размеров и ростом капиталоемкости их производства. В первую очередь, необходимо выделить проблемы отвода тепла, выделяемого из-за непрерывно повышающейся температуры по мере повышения плотности монтажа элементов и частоты, а также возникновения в чипах токов утечки (как из-за работы на высоких частотах, так и за счет проявления квантовых эффектов, в частности, туннельного эффекта [16]).

Эти факторы уже в середине 2000-х гг. стали сдерживать дальнейшее уменьшение напряжения, подаваемое на транзисторы, и особенно рост тактовой (внутренней) частоты. С этого времени закон Деннарда перестал действовать. Показательной является судьба микропроцессора Tejas, разработанного компанией Intel. В 2002–2004 гг. компанией анонсировался выпуск этого процессора с внутренней частотой более 4 ГГц, которую в последующих модификациях предполагалось повысить до

10 ГГц. Интересно отметить, что в тот период Intel планировал обеспечить к 2010 г. производство процессоров с внутренней частотой 15 ГГц, состоящих из 1 млрд транзисторов. Однако в конце 2004 г. разработчикам компании стало ясно, что при заявленных параметрах он характеризуется весьма высоким тепловыделением (впоследствии было установлено, что имеет место степенная зависимость потребляемой процессором мощности от его тактовой частоты с показателем степени, большим 2 [7, 17]). По этой причине в 2005 г. Intel объявил об отказе от производства микропроцессора Tejas, и в 2006 г. перешел к выпуску двухядерного процессора Core, а впоследствии – многоядерных процессоров [7].

Для решения проблемы повышения энергоэффективности микроэлектронных схем (преодоления того, что в англоязычной литературе именуется «power wall») используется широкий спектр методов, начиная с совершенствования применяемых материалов, архитектуры создаваемых приборов, и, кончая использованием программных средств. В начале 2010-х гг. осуществлен переход от производства традиционной планарной к энергоэффективной трехмерной структуре транзисторов (плавниковая конструкция FinFET с топологическим размером 22 и 14 нм, (осм. [7, 18]), достигнуты определенные успехи в разработке кремний-фотонных приборов на одном чипе, в которых проводящие металлические соединения заменяются на оптические (подробнее см. [16]).

Благодаря разработанным научно-техническим решениям, прогресс в области миниатюризации и повышения производительности микроэлектронных устройств и компонентов ЭВМ идет значительно быстрее, чем предполагалось даже в конце 1980-х – начале 1990-х гг. С 1995 по 2015 г. топологический размер полупроводниковых схем уменьшился в 25 раз, причем только за 2005–2015 гг. – почти в 5 раз (см. рис. 2.1).

В 2015 г. Intel начал производство чипов с топологическим размером микросхем 14 нм, причем в 2017 г. предполаглось перейти на топологический размер 10 нм, а впоследствии – на 7 и 5 нм (следует отметить, что другие компании-производители микросхем используют в качестве топологического размера (англ. node) собственные характеристики размеров, которые, несмотря на объявляемую ими малую величину, несопоставимы с более объективной характеристикой топологического размера, основанной на показателе плотности транзисторов в единице площади, к которой близка метрика Intel [7]).

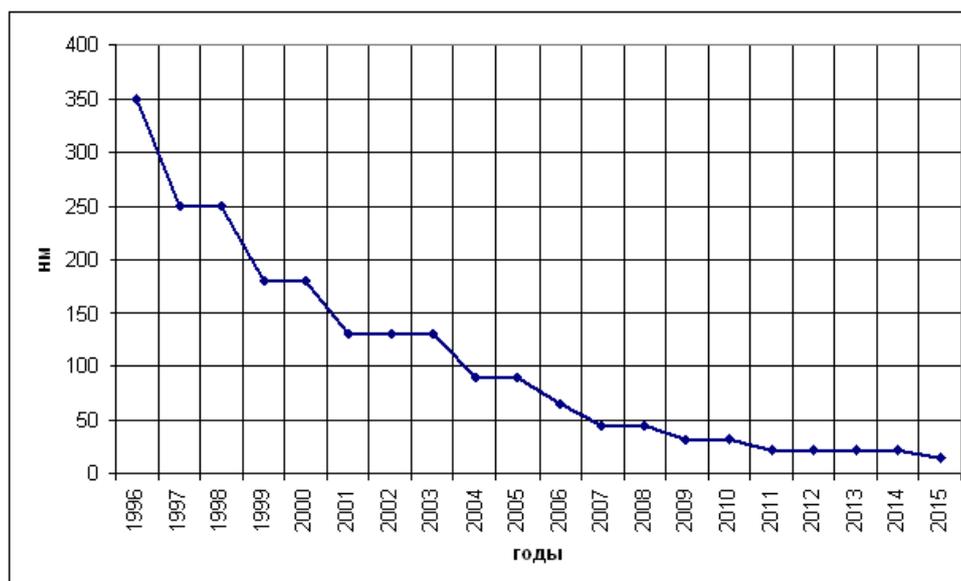


Рис. 2.1. Динамика топологического размера микросхем в компании Intel (построен на основе данных [19, с. 249]; <http://www.intel.com>)

Вероятность перехода компаний-производителей чипов на эти, и еще более совершенные технологические процессы, в последнее время возросла, что, в частности, связано с прогрессом в создании долго ожидавшихся (с начала 2000-х гг.) экономически эффективных установок экстремальной ультрафиолетовой фотолитографии (EUV-сканеров). Так, нидерландская компания ASML Holding NV, являющаяся ведущим производителем установок фотолитографии, уже в 2015 г. производила EUV-сканер NXE:3350B system с оптической мощностью 80 Вт. Однако для рентабельного производства чипов (1500 пластин в сутки) необходимо было повысить выходную мощность излучения (по оценкам этой компании, она должна составлять 200–250 Вт), и решить ряд других проблем (связанных, в частности, с созданием защитных термостойких мембран для фотошаблонов). К 2016 г. на EUV-сканерах компании удалось довести выпуск продукции до 1000 пластин в день, а на установках NXE:3350B system – производительности 1368 пластин в течение 24 часов. При этом коэффициент готовности (availability) созданных EUV-сканеров составлял 70–75%, что было существенно ниже, чем при использовании применяемой иммерсионной технологии (95%). Однако в настоящее время эти проблемы в общем решены и начато производство EUV-сканеров (они уже используются в компании TSMC) [7].

В связи с задержками в сроках начала массового использования EUV-систем, компания Intel планирует освоить производство чипов с топологическими размерами 10 и 7 нм, основываясь на собственных разработках (используя усовершенствованный вариант иммерсионной технологии – multiple patterning). Вместе с тем,

компания TSMC уже использует EUV-установки при освоении технологии с уровнем 5 нм [7].

Наряду с проблемой повышения энергоэффективности, необходимо также отметить проблему роста наукоемкости и капиталоемкости производства, которая обостряется по мере уменьшения топологического размера интегральных схем, характеризующего приближение к физическим пределам роста производительности схем. Так, как уже отмечалось в главе 1, в последнее время ежегодные затраты на исследования и разработки (R&D) только в компании Intel составляют свыше 10 млрд долл (т.е. около 20% от уровня доходов). Такой же высокий уровень характерен для инвестиций компании в основной капитал. Характерно, что даже в кризисные годы объем затрат на исследования и разработки в ведущих компаниях отрасли практически не уменьшается. Достаточно типичной является динамика этих затрат, приведенная на рис. 2.2 по данным компании Intel [7].

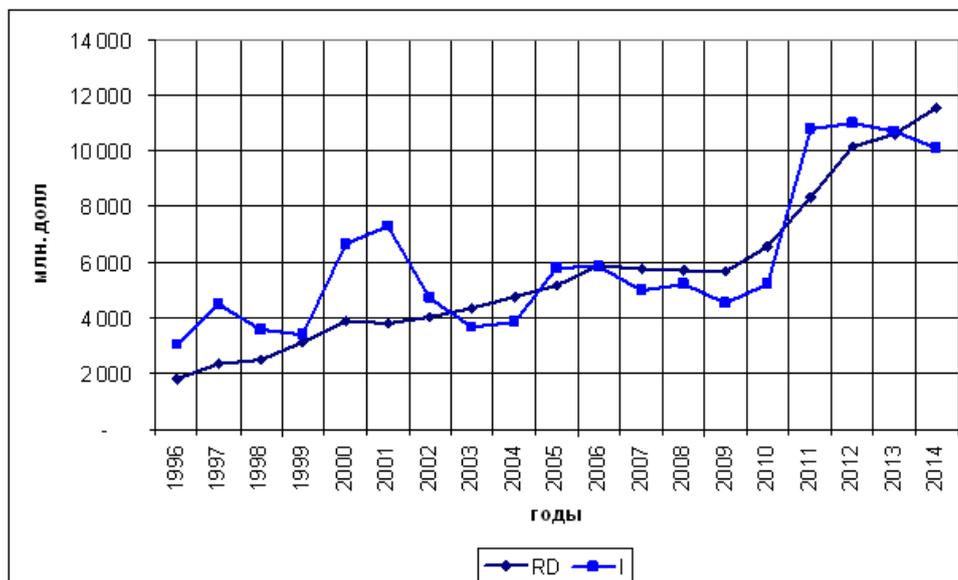


Рис. 2.2. Динамика затрат на исследования и разработки (RD) и капитальных вложений (I) в компании Intel (в млн долл, в текущих ценах: составлено по данным годовых отчетов из www.intel.com)

О значительной величине производственных инвестиций в обновление производства чипов свидетельствует, в частности, и тот факт, что еще в середине 2010-х гг. капитальные вложения Intel в сооружение одной фабрики по производству чипов с топологическим размером схем 14 нм в штате Аризона оценивались в 5 млрд долл. В другой компании, Global Foundries, специализирующейся на производстве чипов, суммарные за время существования производственные инвестиции в фабрику № 1 в г. Дрездене (месячной производственной мощностью 80 тыс. пластин диаметром 300 мм с топологическим размером схем 45 нм) составили 7 млрд долл. Объем

этих инвестиций является одним из наиболее крупных среди зарубежных инвестиций, когда-либо сделанных в Германии. Наглядной характеристикой капиталоемкости производства в отрасли является стоимость EUV-сканера NXE:3350B, которая составляла 75 млн долл., что практически совпадало с ценой современного магистрального узкофюзеляжного самолета Boeing-737-700 (на крупной фабрике требуется не менее 10 таких установок).

Усложнение условий развития заставило компанию Intel отказаться от двухгодичного цикла в используемой с 2007 г. модели технологических изменений (tick-tock) и перейти к трехгодичному циклу (освоение нового технологического процесса – выпуск микропроцессора с новой архитектурой – повышение производительности и энергоэффективности процессоров) [7].

По некоторым оценкам, из-за роста затрат (для схем с размером 5 нм производственные затраты могут увеличиться в 2,5–3 раза по сравнению с затратами на производство схем с размером 14–16 нм!), лишь немногие компании смогут выйти на производство микросхем с топологическим размером 5 нм и менее [20].

Здесь следует отметить, что еще в 2015 г. участники Международной технологической дорожной карты по полупроводникам (ITRS) сочли неэкономичным уменьшение топологического размера микросхем ниже уровня 5 нм после 2021 г. (даже этот размер ненамного больше размера атома кремния, который составляет 0,546 нм!). Дальнейшие методы повышения плотности элементов в чипах связываются с переходом к трехмерным структурам, многослойности и др. Драйверами дизайна чипов, по мнению разработчиков дорожной карты, должны стать такие области применения, как центры обработки данных (ЦОД), Интернет вещей (IoT) и мобильные устройства [21]. Необходимо подчеркнуть, что после снижения спроса на персональные компьютеры в середине 2010-х гг., именно эти направления признаются в настоящее время разработчиками и производителями микросхем и приборов жизненно важными для продолжения своей деятельности.

В последние годы под влиянием перечисленных выше приложений, а также таких приложений, как глубинное обучение (deep learning), наметился переход от микропроцессоров общего назначения к специализированным схемам типа FPGA (Field-Programmable Gate Array – программируемая пользователем вентильная матрица), ASIC (specific integrated circuit– интегральная схема специального назначения), а также к графическим процессорам (GPU). Специализированные схемы типа FPGA и ASIC могут быть эффективно использованы при решении многих задач в области искусственного интеллекта (ИИ), так как обеспечивают более высокое быстродействие, большую энергоэкономичность при меньших затратах (они, в частности, могут быть реализованы на микросхемах с большим топологическим размером).

За счет этого открывается возможность проявления инженерного творчества при создании различных приборов и устройств (например, медицинских и навигационных приборов, устройств мобильной связи и др.), причем не только в крупных компаниях, но и в малом бизнесе [22, 23]. В этом отношении неслучайным является приобретение компанией-лидером Intel компанией Altera, являвшейся крупнейшим производителем FPGA.

Существенное повышение быстродействия и энергоэффективности продемонстрировано и при применении графических процессоров как в качестве ускорителей в суперкомпьютерах, так и при решении различных задач в области ИИ (в частности, при проведении машинного обучения). Так, только за 5 лет быстродействие GPU компании Nvidia повысилось в 25 раз, что значительно (в 2,5 раза) выше, чем могло бы быть при простом следовании, так называемому, закону Мура [24].

Вместе с тем, по некоторым оценкам, дальнейший успех в области использования такого направления ИИ, как машинное обучение, связан не с алгоритмическим обеспечением, а с логистикой, характеризующей данные, т.е. с их качеством, стандартизацией, доступностью [25].

На решение задач в области ИИ (машинное обучение, большие данные, аналитика) ориентируются и разработчики суперкомпьютеров эксафлопсной производительности [26].

Важным драйвером для дальнейшего производства полупроводниковых микросхем и, в частности, микропроцессоров, являются и новейшие коммуникационные технологии. Так, в среднесрочной перспективе во всем мире и у нас предполагается масштабный переход к системе беспроводной мобильной связи 5G. В программе «Цифровая экономика Российской Федерации» во всех крупных городах (1 млн человек и более) намечается обеспечить устойчивое покрытие системой 5G и выше [27].

В целом, можно констатировать, что именно прогресс в создании аппаратных средств ИКТ, и в особенности в области элементной базы, явился главным фактором, способствовавшим быстрому развитию как высокопроизводительных вычислений, так и новых областей искусственного интеллекта и средств связи.

2.3. Экономико-математический анализ и моделирование развития ведущих компаний микроэлектронной промышленности

2.3.1. Влияние топологического развития схем на показатели эффективности производства

Как отмечалось выше, одно из главных направлений увеличения производительности микроэлектронных устройств состоит в уменьшении размеров и повышении степени интеграции схем. Этим объясняется стремление ведущих разработчиков и производителей микроэлектронных изделий уменьшать топологические размеры полупроводниковых схем.

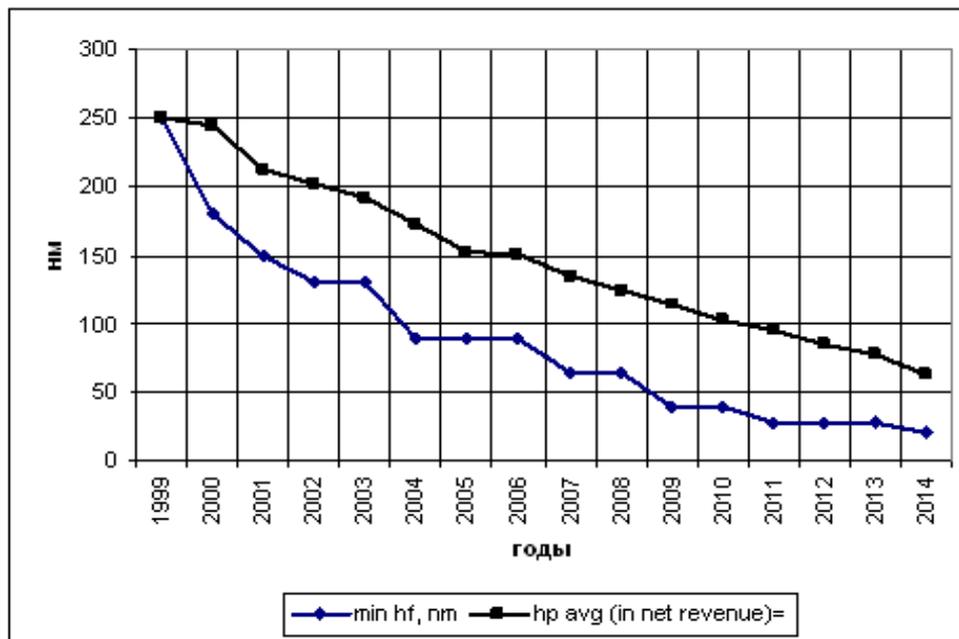


Рис. 2.3. Динамика топологических размеров интегральных схем (минимального $hp_{\min t}$ и среднего $hp_{avg t}$) в компании TSMC (составлено по данным годовых отчетов из www.tsmc.com)

В связи с уменьшением топологического размера схем, производимых во вновь вводимых цехах и фабриках, в ведущих компаниях-производителях происходит снижение средней величины этого показателя (рис. 2.3). Так, в компании TSMC зависимость между минимальным $hp_{\min t}$ и средним (по получаемым доходам) $hp_{avg t}$ топологическими размерами интегральных схем может быть описана следующим регрессионным соотношением, построенным по временным рядам за 1999–2014 гг.:

$$hp_{avg t} = 0,830 hp_{avg t-1} + 0,170 hp_{\min t}, \quad R^2 = 0,999, \quad (2.1)$$

(0,030) (0,046)

$$Q_{Intel t} = \frac{\chi_t z}{(z - 0,801)^2} I_{Intel t}, \text{ где } \chi_t = \begin{cases} 3,627 & \text{при } t \leq 2009; \\ 1,283 & \text{при } t > 2009, \end{cases} \quad (2.3)$$

где z – оператор сдвига: $zx_t = x_{t+1}$ (коэффициент вариации зависимости составляет $v = 0,116$). Более низкая отдача от инвестиций после 2009 г. (характеризуется значением параметра χ_t) связана со стабилизацией спроса на персональные компьютеры (PC) и ростом капитальных затрат ввиду перехода на более совершенные и капиталоемкие технологические процессы (32, 22 и 14 нм).

Для компании AMD, которая до 2009 г. также относилась к *вертикально* интегрированным компаниям, на основе временных рядов из годовых отчетов за 1997–2009 гг. (см. www.amd.com) получена следующая операторная зависимость между объемами производства микропроцессоров x86 $Q_{AMD t}$ (в млн ед.) и инвестиций в основной капитал $I_{AMD t}$ (в млрд. долл):

$$Q_{AMD t} = \frac{15,6z}{(z - 0,494)^2} I_{AMD t}. \quad (2.4)$$

Об удовлетворительном качестве полученной зависимости свидетельствует приемлемое значение коэффициента вариации ($v = 0,092$). Полученная зависимость позволяет рассчитать среднюю продолжительность выпуска изделий за счет осуществленных инвестиций T_L и $T_{0,01}$ – время выхода на уровни производства, отличающиеся не более 1% от установившихся, при постоянных уровнях инвестиций. Значения этих показателей составляют соответственно $T_L = \frac{2 \cdot 0,494}{1 - 0,494} \approx 2$ года и $T_{0,01} = 9$ лет.

Меньшая инерционность связи между исследуемыми показателями в компании AMD (динамика связей характеризуется числами в знаменателях соотношений (2.3) и (2.4)) обусловлена тем, что эта компания, являющаяся «вторым источником» производства микропроцессоров x86 для персональных компьютеров архитектуры IBM, значительную часть продукции выпускала по лицензиям компании Intel [7].

2.4. Прогнозирование динамики показателей рынка микропроцессоров x86 на среднесрочный период

2.4.1. Используемая модель

Построенные зависимости используются далее в качестве составных блоков агрегированных моделей прогнозирования динамики показателей рынка

микропроцессоров x86, которые представляют собой один из вариантов динамической модели рационального поведения участников олигополии в виде линейной динамической игры по Нэшу–Курно [28, 29]. Предполагается, что олигополия состоит из N фирм. Центральным блоком модели является следующая зависимость, связывающая объемы производства Q_{it} со входной переменной u_{it} (производственными инвестициями или вводом мощностей), i – индекс фирмы, $i = 1, 2, \dots, N$:

$$Q_{it} = W_i(z) u_{it} + Q_{0it} = \frac{B_i(z)}{A_i(z)} u_{it} + Q_{0it}, \quad (2.5)$$

где $W_i(z) = B_i(z) / A_i(z)$ – передаточная функция, причем $A_i(z)$, $B_i(z)$ – полиномы относительно переменной z , представляющей собой оператор сдвига: $zx_t = x_{t+1}$

$$A_i(z) = \sum_{k=0}^n a_{ik} z^k, \quad B_i(z) = \sum_{j=0}^m b_{ij} z^j, \quad m \leq n, \quad (2.6)$$

имеющие корни, не превышающие по модулю 1, причем $W_i(1) > 0$; Q_{0it} – объем производства при отсутствии инвестиций.

Другой блок модели – обратная функция спроса. В модели предполагается баланс суммарного спроса D_t и предложения Q_t , т.е. $D_t = Q_t = \sum_{i=1}^N Q_{it}$ и линейная зависимость цены на рынке p_t от объема спроса:

$$p_t = a - bD_t = a - bQ_t, \quad (2.7)$$

a , b – постоянные параметры.

Предполагается, что олигополисты максимизируют чистую текущую стоимость (NPV) с учетом затрат регулирования (adjustment costs):

$$J_i = \sum_{t=0}^{\infty} \beta^t \left[(p_t - c_i) Q_{it} - q_i u_{it} - \frac{1}{2} \rho_i u_{it}^2 \right] \rightarrow \max_{u_{it}}, \quad (2.8)$$

где $\beta = 1 / (1 + r)$ – дисконтирующий множитель, соответствующий ставке дисконтирования r ; p_t – цена продукции; c_i – средние удельные производственные издержки; q_i – стоимость единицы мощностей; $\frac{1}{2} \rho_i u_{it}^2$ – затраты регулирования (adjustment cost, см., например, [30]), причем $\rho_i > 0$ – коэффициент, характеризующий инвестиционные возможности олигополистов (при прочих равных условиях, чем меньше его величина, тем эти возможности больше), $i = 1, 2, \dots, N$. Управляющими переменными в модели являются вводы мощностей (или инвестиции в основной капитал) u_{it} , $i = 1, 2, \dots, N$.

2.4.2. Оптимальные по Нэшу–Курно стратегии участников рынка

Для проведения прогнозных расчетов на основе модели (2.5)–(2.8) могут быть использованы 2 подхода: 1). основанный на представлении модели в пространстве состояний и использовании обобщенных (generalized, в англоязычной литературе используются также термин coupled) матричных уравнений Риккати (см., например, [31, 32]; 2). основанный на использовании операционного исчисления [28, 29]. Эти подходы взаимно дополняют друг друга.

При использовании первого подхода для расчета как разомкнутых (open-loop), так и замкнутых (feedback или Markov perfect) оптимальных по Нэшу–Курно стратегий следует предварительно представить модель (2.5)–(2.8) в эквивалентной форме в пространстве состояний:

$$X_t = AX_{t-1} + \sum_{i=1}^N B_i u_{it} + D\xi_t, \quad (2.9)$$

$$J_i = \sum_{t=0}^{\infty} \beta^t \left(\frac{1}{2} X_t' H_i X_t - C_{0i}' X_t - q_{it}' u_{it} - \frac{1}{2} \rho_i u_{it}^2 \right) \rightarrow \max_{u_{it}}, \quad (2.10)$$

где матрицы и векторы $A, B_i, D, H_i, C_{0i}, q_{it}, X_t, \xi_t, i = 1, 2, \dots, N$ связаны с параметрами и переменными исходной модели. Получаемые оптимальные стратегии участников олигополии u_{it} линейно связаны с вектором состояния системы (2.9):

$$u_{it} = K_{it} X_{t-1} + \eta_{it}, \quad (2.11)$$

в котором K_{it} и η_{it} – векторы, зависят от решений обобщенных уравнений Риккати (Basar and Olsder, 1995).

При определении оптимальных разомкнутых (open-loop) игровых стратегий олигополистов с критерием (2.8), более наглядным и удобным для экономического анализа является второй подход, основанный на использовании операционного исчисления (широко распространенного при исследовании дискретных систем управления и в теории связи Z-преобразования, см., например, [33, 34]) и нахождении экстремума функционалов в гильбертовом пространстве (см. [28, 29]). Используя предложенный метод, из необходимого условия экстремума функционала (2.8) можно получить формулы для расчета оптимального управления u_{it} (производственных инвестиций и др.) и объемов производства Q_{it} i -го олиголиста, максимизирующего критерий NPV с учетом затрат регулирования [29]:

$$Q_{it} = W_i(z) u_{it} = \frac{\Gamma_i(z, (\beta z)^{-1})}{b} (p_t - PL_i - bQ_{0it}), \quad (2.12)$$

где $PL_i = c_i + q_i / W(1 + r_i)$ – лимитирующие (приведенные) затраты i -й фирмы;

$$\Gamma_i[z, (\beta z)^{-1}] = \frac{bW_i(z)W_i((\beta z)^{-1})}{\rho_i + bW_i(z)W_i((\beta z)^{-1})}, \quad i = 1, 2, \dots, N; \quad (2.13)$$

$$p_i = \frac{1}{1 + \sum_{i=1}^N \Gamma_i[z, (\beta z)^{-1}]} \left\{ a + \sum_{i=1}^N \Gamma_i[z, (\beta z)^{-1}] (PL_i + bQ_{0it}) \right\}. \quad (2.14)$$

Для получения выражений, характеризующих динамику рассматриваемых показателей во временной области, необходимо провести факторизацию правых частей выражений (2.13)–(2.14). Факторизация многочленов может быть проведена, например, в среде MATLAB с использованием процедур поиска корней полиномов, а также процедур формирования соединений систем, реализованных в Control Systems Toolbox (таких, в частности, как `series(.)`, `parallel(.)`, `feedback(.)` и др. При невысоком порядке передаточной функции $W_i(z)$ такие расчеты могут быть также проведены с использованием нашедших широкое распространение электронных таблиц [35].

Используя свойства Z-преобразования, можно при постоянных PL_i , основываясь на (2.12)–(2.14), оценить установившиеся уровни производства участников рынка $Q_{i\infty}$ [29]:

$$Q_{i\infty} = \frac{\Gamma_i(1, (1+r))}{b} (p_\infty - PL_i), \quad (2.12a)$$

цены $p_\infty = a - b \sum_{i=1}^N Q_{i\infty}$, рыночных долей фирм $MS_i(\infty)$, а также других показателей.

Из (2.12)–(2.14), (2.12a) следует, что объемы производства и рыночная доля фирм тем выше, чем меньше их лимитирующие затраты PL_i и коэффициенты ρ_i , характеризующие их инвестиционные возможности (их значение влияет на величину $\Gamma_i(1, 1+r)$). Последнее соотношение может быть использовано при оценке значений коэффициентов ρ_i , при которых достигаются желаемые (целевые) установившиеся уровни производства участников рынка $Q_{i\infty}$. Так, очевидно:

$$\rho_i = \left[\frac{(p_\infty - PL_i)}{Q_{i\infty}} - b \right] W_i(1)W_i(1+r) = (\delta_i - 1) b W_i(1)W_i(1+r), \quad (2.15)$$

где $\delta_i = \frac{p_\infty - PL_i}{bQ_{i\infty}}$ – характеристика инвестиционных ограничений i -й фирмы (в статической модели Курно, в которой отсутствуют такие ограничения, $\delta_i = 1$) $i = 1, 2, \dots, N$. После нахождения коэффициентов ρ_i может быть рассчитана динамика объемов производства олигополистов Q_{it} , цены и другие рыночные показатели. С другой стороны, на основе соотношения (2.15) можно приближенно оценить

фактические значения коэффициентов p_i . Для этого следует заменить $Q_{i\infty}$ и p_{∞} на среднегодовые значения этих показателей за представительный период: $Q_{i\text{avg}}$ и p_{avg} .

2.4.3. Прогнозные сценарии динамики показателей рынка микропроцессоров x86

Как известно, рынок микропроцессоров архитектуры x86 представляет собой дуополию, участниками которой являются компании Intel и AMD с рыночными долями соответственно свыше 80 и менее 20% поставок и продаж. В связи с этим, сценарии развития данного рынка могут быть построены путем решения игровой задачи (2.12)–(2.15) при разных гипотезах изменения параметров моделей.

В настоящей работе в качестве блоков, описывающих связи между производственными инвестициями и объемами производства микропроцессоров в компаниях использованы несколько измененные зависимости (2.3) и (2.4). Так, в расчетах принято, что значение χ_t в зависимости (2.3) равно среднеарифметическому полученных значений до и после 2009 г., т.е. $\chi_t = (3,627 + 1,283) / 2 = 2,455$. Несмотря на то, что с 2009 г. чипы с микропроцессорами для компании AMD производятся в основном на фабриках компании Global Foundries, динамика связи между производственными инвестициями в развитие фабрик и объемами производства микропроцессоров в компании AMD может, по нашему мнению, сохраниться и в среднесрочной перспективе.

В результате эконометрического анализа обратной функции спроса (2.7) с начала 2000-х до 2014 г. (исследовалось 5 линейных регрессионных зависимостей для 5 периодов времени) получены стабильные значения параметров a и b с высокими значениями t -статистик (для параметра a : 7–10, а для параметра b – 2–3) при приемлемых значениях коэффициента детерминации R^2 (порядка 0,5). В итоге в расчетах в качестве базовой была использована следующая обратная функция спроса на микропроцессоры архитектуры x86 (включая микропроцессоры для настольных персональных компьютеров, ноутбуков и серверов):

$$p_t = 180,285 - 0,159 Q_{total\ t}, \quad (2.16)$$

(23,699) (0,081)

где $Q_{total\ t} = Q_{Intel\ t} + Q_{AMD\ t}$ – суммарный объем поставок микропроцессоров архитектуры x86 в млн ед., p_t – средняя цена микропроцессоров компании Intel в долл/ед. Средняя цена микропроцессоров компании AMD составляла примерно 0,6 от соответствующего показателя компании Intel (по нашим оценкам эта величина составляла в 2010–2014 гг. 0,584).

Основными потребителями микропроцессоров x86 являются производители персональных компьютеров, которые с начала 2010-х гг. испытывают стабилизацию (с 2010 по 2012 г.) и даже падение спроса на РС (с 2013 г., рис. 2.4). По мнению аналитиков, причины этого состояли в том, что возможности персональных компьютеров с последними поколениями микропроцессоров x86 превосходят потребности большинства потребителей, которых вполне устраивают возможности эксплуатируемых компьютеров и мобильных устройств (смартфонов и планшетов). Другой фактор обусловлен ухудшением общеэкономической ситуации в мире. В связи с этими факторами произошло увеличение продолжительности эксплуатации РС, что в итоге отразилось на спросе как на них самих, так и на комплектующие для них [7].



Рис. 2.4. Динамика общемировых продаж РС (построено по данным в Wikipedia)

Ввиду отсутствия надежных данных о наличном парке РС, при формировании сценариев изменения показателей рынка микропроцессоров x86 приходится задаваться гипотезами о тех или иных изменениях параметров обратной функции спроса (2.23). В работе [7] в результате решения игровой задачи (2.5)–(2.8) построены разомкнутые (open-loop) оптимальные по Нэшу–Курно стратегии участников дуополии при разных значениях параметров функции спроса ($a = 160,2; 170,2; 180,2; b = 0,149; 0,159; 0,169; 0,179$) и соответствующие им ключевые показатели рынка микропроцессоров x86. Результаты расчетов приведены на рис. 2.5–2.7.

Для большинства сценариев проведенные расчеты показывают стабилизацию или уменьшение объемов суммарного производства рассматриваемых процессоров $Q_{total t}$ (максимальное уменьшение может составить около 15% к 2021 г. относительно 2015 г.). При уменьшении параметра b до уровня $b = 0,149$ возможен рост производства микропроцессоров относительно 2015 г. на 10%, а к 2025 г. – на 14%.

Такой сценарий может произойти, в частности, в случае более интенсивного, чем в середине 2010-х гг., выбытия эксплуатируемых компьютеров, и при улучшении общемировой экономической ситуации (рис. 2.7).

Средняя цена микропроцессоров x86 p_t при всех сценариях сценариев может быть ниже, чем в 2015 г., а рыночная доля компании AMD в большинстве сценариев может быть несколько выше, чем в базовом 2015 г. (рис. 2.6–2.7).

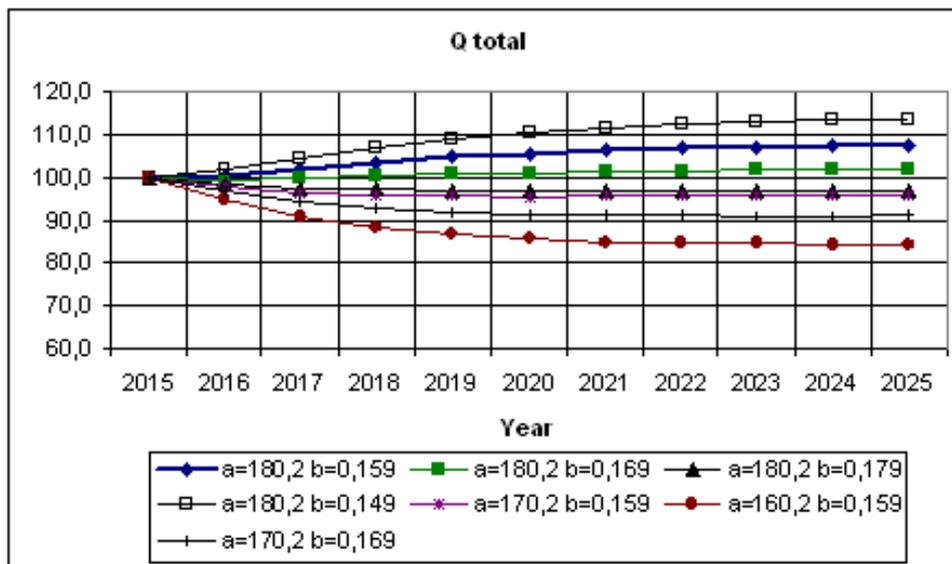


Рис. 2.5. Динамика суммарных объемов производства микропроцессоров x86 $Q_{tot t}$ при разных значениях параметров обратной функции спроса (2015 = 100)

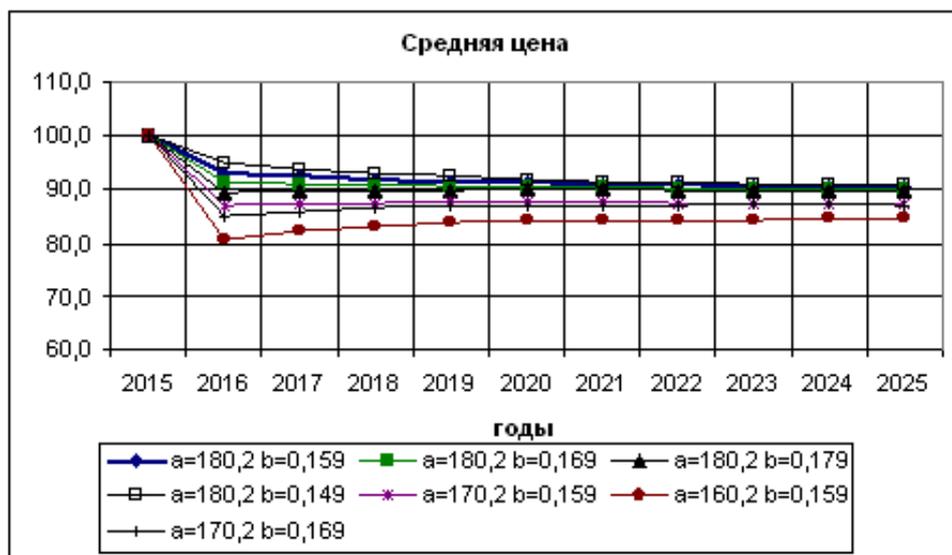


Рис. 2.6. Динамика средней цены микропроцессоров x86 p_t при разных значениях параметров обратной функции спроса (2015 = 100)

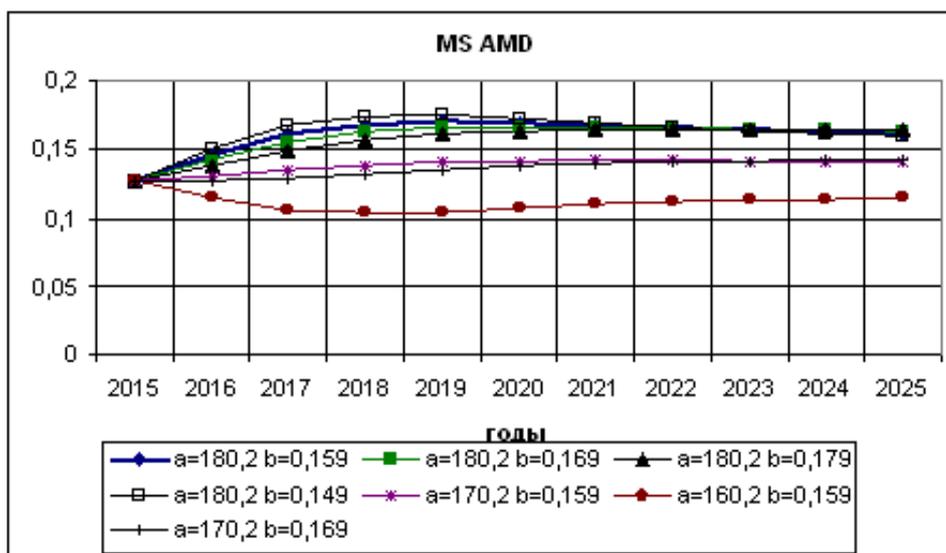


Рис. 2.7. Динамика рыночной доли (MS) компании AMD в производстве микропроцессоров x86 при разных значениях параметров обратной функции спроса (2015 = 100)

В связи с насыщением в середине 2010-х гг. рынка персональных компьютеров, ведущие производители микропроцессоров стали переключаться на выпуск более перспективной продукции, с которой связываются повышенные темпы роста спроса. Такими видами продукции для Intel являются, в частности микропроцессоры и другие чипы для серверов (для центров обработки данных – ЦОД), для Интернета вещей (IoT), для робототехнических устройств. Эти направления уже в 2016 г. обеспечивали почти 40% доходов и более 70% операционной прибыли компании. Следует отметить и то, что, если средняя цена производимого в компании Intel микропроцессора x86 составляла порядка 140 долл., то средняя цена микропроцессора для серверов – свыше 600 долл. Цена на наиболее мощные многоядерные микропроцессоры этой компании (например, на 18-ядерный процессор Intel Xeon E7-8890 v3) превышала 7000 долл. В связи с наличием такой разницы, все больше компаний стремится выйти на высокорентабельный и быстрорастущий рынок микропроцессоров для серверов.

2.5. Исследование перспективной структуры рынка микропроцессоров для серверов (с архитектурой x86 и ARM)

В настоящее время в производимых серверах используются микропроцессоры с архитектурой x86 и RISC. Доля первого типа серверов в объеме продаж (в млрд долл) в середине 2010-х гг., составляла 70%, а в объеме поставок (в млн ед.)

приближается к 90%. С этого времени главный производитель микропроцессоров с архитектурой x86, компания Intel (на нее приходится 99% поставок процессоров для серверов, причем практически все они относились к серии Intel Xeon), начинает испытывать конкуренцию с выходящими на рынок серверов производителями микропроцессоров с архитектурой ARM (она также представляет разновидность архитектуры RISC).

По мнению специалистов, процессоры с архитектурой ARM в перспективе могут обеспечить большую экономичность за счет существенно меньшей энергоемкости, что подтверждается стремительным распространением и практически полной монополизацией микропроцессорами с архитектурой ARM рынка мобильных устройств. По некоторым оценкам, в 2020 г. доля серверов, оснащенных микропроцессорами с этой архитектурой, в общем производстве микропроцессоров x86+ARM могла достигнуть 20–25% (<http://www.computerworld.in/news/intel-faces-challenge-server-market-new-arm-chips>). Однако, по мнению аналитиков, массовое производство такой продукции компаниями-лицензиарами ARM-архитектуры должно было начаться в 2017–2018 гг. (табл. 2.1). Появление серверов с ARM-архитектурой приветствуется провайдерами Интернета и облачных сервисов, а также в финансовом секторе (финансовый гигант Morgan Stanley остался удовлетворенным результатами тестирования сервера с ARM-микропроцессором X-Gene компании AppliedMicro). Этот сегмент потребителей заинтересован в снижении зависимости от одного монополиста микропроцессоров x86.

Таблица 2.1.

Некоторые разработчики и возможные производители серверных микропроцессоров архитектуры ARM

| Компания-лицензиар ARM-архитектуры | Микропроцессор | Возможный производитель чипов, техпроцесс | Год начала производства |
|------------------------------------|----------------|---|-------------------------|
| Applied Micro | X-Gene 3 | TSMC, 16 nm FinFET | Конец 2016 |
| Cavium | ThunderX2 | Global Foundries, 14 nm FinFET | Первая половина 2017 г. |
| Qualcomm | QCOM | Samsung, 14 nm process | Не ранее 2017 г. |
| AMD | K-12 | н.д. | 2017 |
| Applied Micro | X-Gene 3 | TSMC, 16 nm FinFET | Конец 2016 |

Составлено на основе данных в: URL: <http://seekingalpha.com/article/3983132-intel-arm-servers-renew-attack>

К грядущим изменениям на рассматриваемом рынке серьезно готовится и Intel. Так, компанией ведутся работы по созданию гибридного чипа, состоящего из микропроцессора Xeon и FPGA. По результатам исследований в компаниях Intel и Google, использование FPGA в качестве ускорителя оказывалось эффективнее, чем графического процессора Tesla 20 GPU компании Nvidia с точки зрения, как

быстродействия, так и энергоэффективности. В Intel считали, что к 2020 г. до 1/3 провайдеров облачных сервисов будет использовать такие гибридные чипы.

С целью исследования сценариев развития рынка, в [7] на основе данных компаний IDC и Gartner построена следующая обратная функция спроса на микропроцессоры для серверов архитектуры x86:

$$p_t = 697,684 + 24,087 R \& D_t - 16,351 Q_{x86t}, R^2 = 0,802, \quad (2.17)$$

(168,787) (4,888) (8,859)

где Q_{x86t} – объем поставок микропроцессоров архитектуры x86 для серверов в млн ед.; p_t – средняя цена этих микропроцессоров в долл./ед.; $R \& D_t$ – объем исследований и разработок в компании Intel в млрд долл.

Прогнозный уровень исследований и разработок в компании Intel рассчитывался на основе экономико-статистической зависимости (2.2), связывающей $R \& D_t$ в млн долл с минимальным топологическим размером интегральных схем $hp_{\min t}$ (см. п. 2.3), причем перспективные значения показателя $hp_{\min t}$ брались из дорожных карт компании.

При проведении расчетов предполагалось, что рынок микропроцессоров архитектуры x86 и ARM для серверов представляет собой олигополию, участниками которой являются с одной стороны компания Intel (практически все производство микропроцессоров архитектуры x86 будет сосредоточено на ее фабриках), а с другой стороны – компании, выпускающие микропроцессоры архитектуры ARM.

Предполагалось, что аргументами обратной функции спроса (24), наряду с $R \& D_t$, являются $Q_{total t} = Q_{1t} + Q_{2t}$, где индекс 1 относится к компаниям, производящей микропроцессоры архитектуры ARM, а индекс 2 – к компании Intel, производящей микропроцессоры архитектуры x86.

Кроме того, предполагалось, что технологический уровень оборудования компаний, выпускающих микропроцессоры архитектуры ARM одинаков. Полученные в результате эконометрического анализа операторные зависимости между инвестициями в производство микропроцессоров для серверов I_{it} (в млрд долл.) и объемами производства микропроцессоров для серверов двух типов Q_{it} (в млн ед.), а также удельные затраты на производство микропроцессоров для базового варианта (в долл) приведены в табл. 2.2.

Таблица 2.2

Показатели компаний для базового сценария

| Компания | Передаточная функция | Удельные затраты на производство микропроцессора, c_1, c_2 , долл. | Коэффициенты при затратах регулирования ρ_1, ρ_2 |
|--|---|--|--|
| Производители микропроцессоров архитектуры ARM | $Q_{1t} = \frac{3,068 z}{(z - 0,494)^2} I_{1t}$ | 160 | 6299 ¹ |
| Intel | $Q_{2t} = \frac{0,569 z}{(z - 0,822)^2} I_{2t}$ | 160 | 363 ² |

¹ Рассчитано в соответствии с (2.15) при среднегодовом значении δ для компании AMD на рынке микропроцессоров.

² Рассчитано в соответствии с (2.15) на основе среднегодовых данных на рынке серверных микропроцессоров.

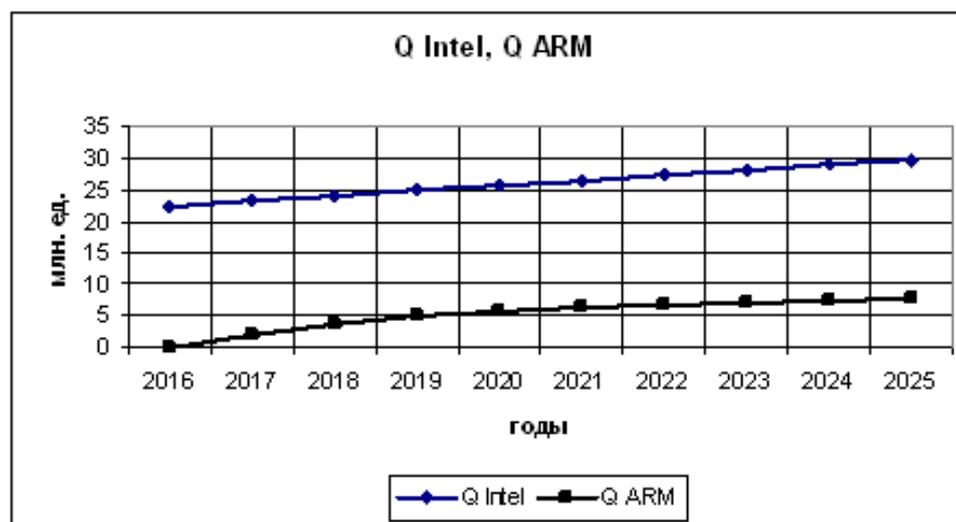


Рис. 2.8. Динамика объемов производства микропроцессоров для серверов архитектуры ARM Q_{1t} и x86 Q_{2t} в базовом сценарии

На рис. 2.9 представлена динамика объемов производства микропроцессоров для серверов архитектуры ARM Q_{1t} и x86 Q_{2t} в базовом сценарии для случая дуополии. При этом сценарии в 2020 г. объем производства микропроцессоров архитектуры ARM должен был превысить 5 млн ед., а архитектуры x86 – 25 млн ед.

Для случая дуополии рассматривались возможные в перспективе варианты структуры рынка, соответствующие различным величинам соотношений между удельными операционными затратами компаний $\eta_c = c_1/c_2$, а также между коэффициентами, характеризующими их инвестиционные возможности компаний $\eta_p = \rho_1/\rho_2$. Некоторые результаты расчетов, проведенных на основе модели (2.5)–(2.8) с использованием расширенных уравнений Риккати приведены на рис. 2.9–2.10.

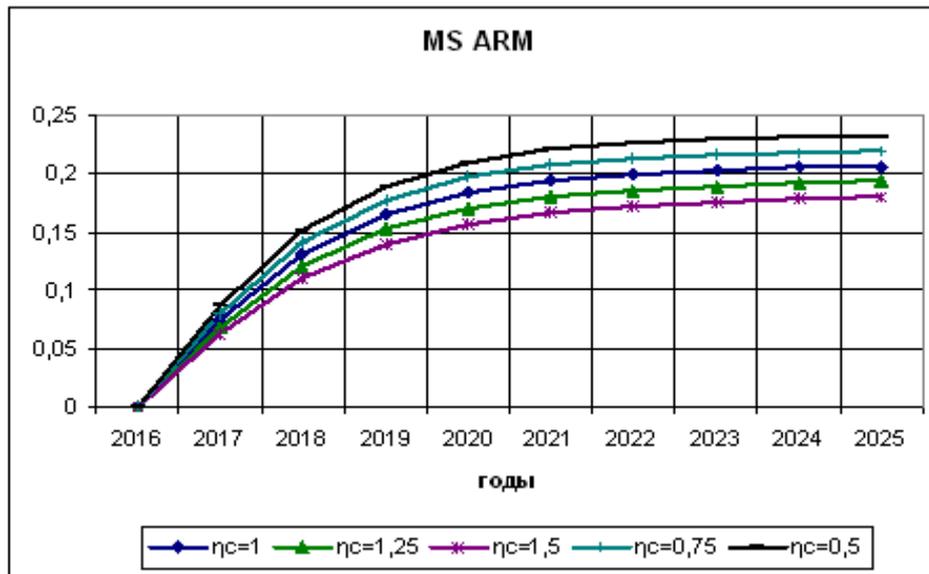


Рис. 2.9. Динамика рыночной доли MS_{1t} микропроцессоров для серверов архитектуры ARM при различной величине соотношений между удельными операционными затратами компаний η_c (при условии $\eta_p = 1$)

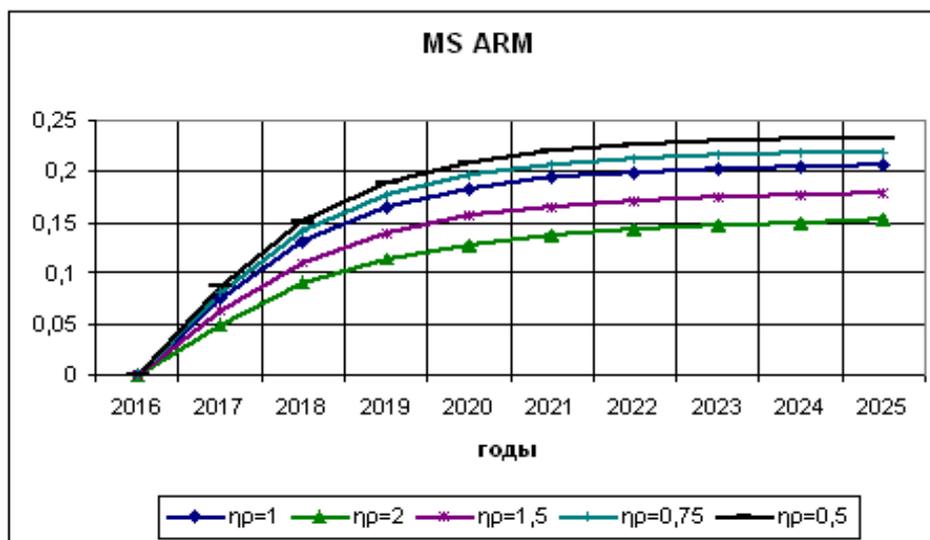


Рис. 2.10. Динамика рыночной доли MS_{1t} микропроцессоров для серверов архитектуры ARM при различной величине соотношений между коэффициентами η_p (при условии $\eta_c = 1$)

На рис. 2.11 представлены варианты динамики рыночной доли MS_{1t} микропроцессоров архитектуры ARM при величинах соотношений между удельными операционными затратами компаний η_c , равных 1,0, 1,25, 1,5, 0,75, и 0,50 (при этом $\eta_p = 1$), а на рис. 2.11 – при величинах соотношений между коэффициентами η_p , равных 1,0, 2,0, 1,5, 0,75, 0,5. При базовом сценарии рыночная доля MS_{1t} микропроцессоров архитектуры ARM должна была бы приблизиться к 20% в 2021 г. Расчеты

показывают, что структура рассматриваемого рынка в перспективе несколько более чувствительна к изменениям величин соотношений между коэффициентами, характеризующими их инвестиционные возможности компаний $\eta_p = \rho_1/\rho_2$, чем к изменениям величин η_c (см. рис. 2.9, 2.10).

Кроме того, исследовалось также влияние числа новых компаний, производящих микропроцессоры архитектуры ARM, на показатели рынка. При этом рассматривалось два случая: 1) инвестиционные возможности каждой из N_1 этих компаний такие же, как и в случае дуополии в базовом сценарии ($\rho_{1i} = \rho_1 = 6299$): 2) суммарные инвестиционные возможности всех этих компаний такие же, как и для одной компании в случае дуополии в базовом сценарии (т.е. $\rho_{1i} = \rho_1 = 6299 \cdot N_1$). На рис. 2.11, 2.12 представлена динамика рыночной доли MS_{1t} серверных микропроцессоров архитектуры ARM для этих двух случаев при числе новых компаний-производителей микропроцессоров ARM, равном 1, 2, 4, 8. По нашим расчетам, расширение числа новых компаний, обладающих значительными инвестиционными возможностями, до 4 и 8, позволило бы уже к 2020 г. повысить рыночную долю микропроцессоров ARM до 45 и 55% (рис. 2.11). При скромных инвестиционных возможностях новых компаний расширение их числа приведет к незначительному эффекту (рис. 2.12). Однако главное условие успешного закрепления на рассматриваемом рынке микропроцессоров архитектуры ARM связано с успешным технологическим прорывом, который может быть осуществлен при решении множества проблем в области программного обеспечения и формирования новой экосистемы (software ecosystem).

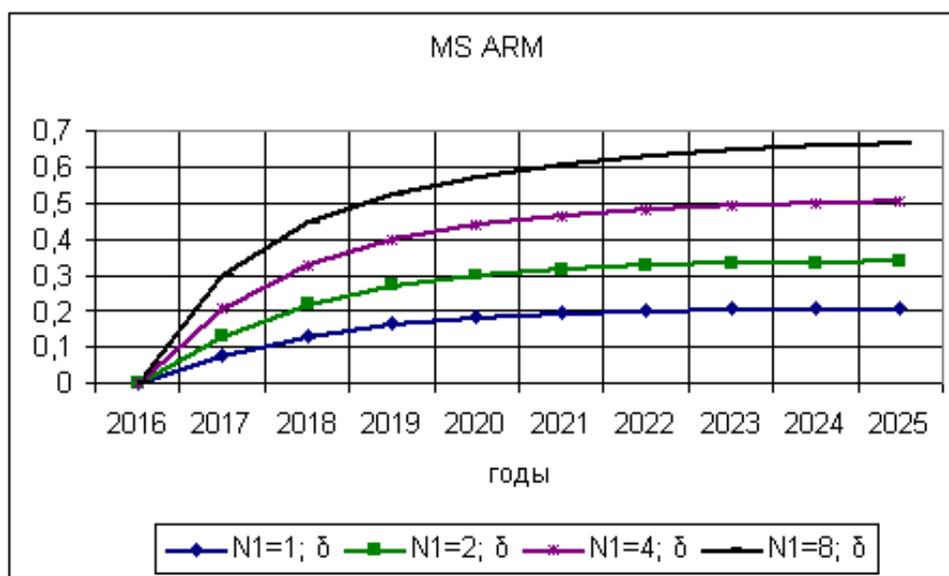


Рис. 2.11. Динамика рыночной доли MS_{1t} микропроцессоров для серверов архитектуры ARM при значительных инвестиционных возможностях каждой из новых компаний при разном их числе N_1

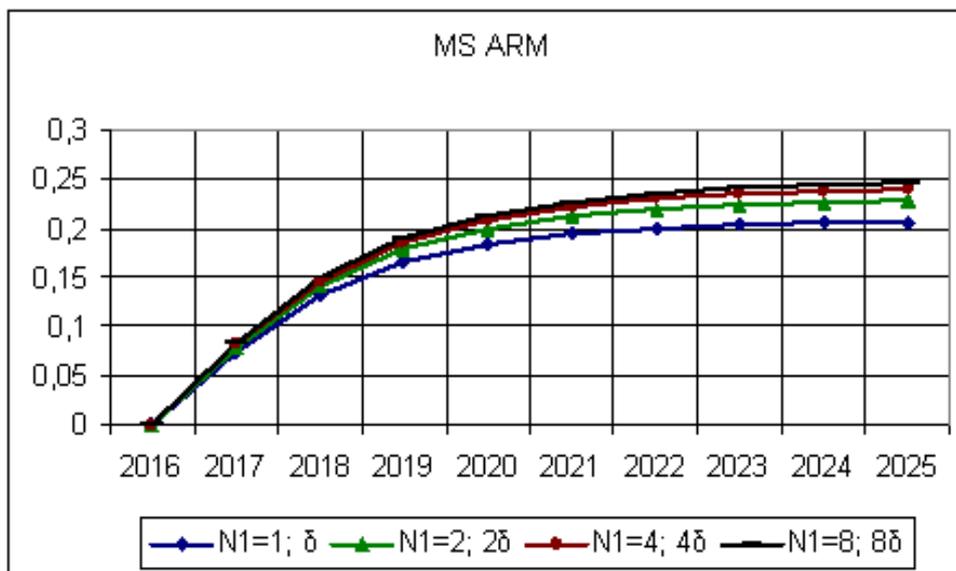


Рис. 2.12. Динамика рыночной доли MS_{1t} микропроцессоров для серверов архитектуры ARM при скромных инвестиционных возможностях каждой из новых компаний при разном их числе N_1

2.6. Возможный подход к формированию рыночных структур в микроэлектронике

При организации новых производств, имеющих жизненно важное значение для экономики и обороноспособности страны, необходимо не допускать монополизацию производства. В связи с этим, за рубежом в ряде отраслей, особенно в электронной промышленности, широко используется идея «второго источника», появление которой связывают с бывшим министром обороны Р. Макнамара [36]. Введение «второго источника» позволяет не только снизить риски прекращения разработки и производства жизненно необходимой продукции, но и, за счет возникновения конкуренции во многих случаях, снизить цену для потребителей. Следует отметить, что в оборонно-промышленном комплексе СССР также поддерживалась конкуренция между предприятиями, хотя и не на рыночной основе.

Классическим примером «второго источника» является компания AMD, которая в 1982 г. «по запросу IBM» подписала «соглашение об осуществлении функции второго поставщика микропроцессоров для персональных компьютеров IBM наряду с Intel» (<http://www.amd.com/ru-ru/who-we-are/corporate-information/history>). Некоторые другие примеры соглашений о «втором источнике» в электронной промышленности приведены в табл. 2.3.

Таблица 2.3

Примеры соглашений о «втором источнике» в электронной промышленности

| Компания – разработчик нового продукта (технологии) | Компания – «второй источник» | Рынок | Лицензируемый продукт | Год заключения соглашения |
|---|------------------------------|---|--|---------------------------|
| Fci electronics | Samtec inc. | Высокоскоростные межсоединения | Высокоскоростные соединения следующего поколения ehamax [®] , разработанные fci electronics | 2014 |
| Soitec | Intelliepi | Полупроводники на основе арсенида галлия (gaas) | Технологии компании soitec | 2013 |
| Integrated device technology, inc. (idt) | Texas instruments inc. | Переключательные устройства для телекоммуникационных систем | Высокоскоростные переключательные устройства, разработанные компанией idt | 2001 |
| Microsemi | Motorola | Выпрямители для смартфонов и ноутбуков на основе эффекта шоттки | Пакет выпрямителей powermite(r) малого размера, разработанный компанией microsemi | 1996 |

Источники: URL: <http://www.fci.com/en/press/press-library/2014/corporate-and-financial/fci-samtec-announces-new-strategic-partnership.html>; <http://www.soitec.com/en/news/press-releases/soitec-and-intelliepi-enter-in-collaboration-agreement-to-better-serve-the-gallium-arsenide-gaas-market-1396/>; http://www.eetimes.com/document.asp?doc_id=1129403.

Рассмотренный в пп. 2.4–2.5 настоящей главы игровой подход может быть также использован и при формировании рыночных структур в высокотехнологичных отраслях, и в частности, в микроэлектронной отрасли страны [8, 9]. Для иллюстрации этого подхода ниже рассматривается условный пример, в котором предполагается, что новое производство микропроцессоров для компьютеров может быть освоено в двух компаниях, (т.е. создается дуополия). Задача государства может состоять в расширении инвестиционных возможностей фирм, ориентированных на достижение максимальной экономической эффективности [8, 9, 37, 38].

Так, из (2.12a) следует, что фирмы могут повысить объемы производства и рыночную долю $MS_1(\infty)$ и $MS_2(\infty)$ за счет уменьшения лимитирующих затрат PL_1 и коэффициентов ρ_i , характеризующих их инвестиционные возможности (их значение влияет на величину $\Gamma_i(1, (1+r))$). Следовательно, желаемая рыночная структура может быть обеспечена регулирующим органом (или ассоциацией потребителей производимого фирмами продукта) как путем предоставления субсидий, так и путем расширения инвестиционных возможностей отстающих участников.

Для реализации последнего способа необходимо рассчитать значения коэффициентов ρ_i , при которых достигаются желаемые установившиеся уровни

производства компаний $Q_{i\infty}$. Основываясь на (2.15), где $p_\infty = a - b(Q_{1\infty} + Q_{2\infty})$, можно определить необходимые значения коэффициентов ρ_i .

После нахождения коэффициентов ρ_i , на основе (2.12)–(2.14) вычисляется динамика объемов производства дуополистов Q_{it} , цены и других рыночных показателей (подробнее см. [8, 9]). На рис. 2.13 и 2.14 приведена динамика объемов производства микропроцессоров и инвестиций компаний в дуополии. В расчетах принято, что передаточные функции, связывающие инвестиции (в млн долл.) и объемы производства процессоров (в млн ед.) имеют следующий вид:

$$Q_{it} = \frac{0,02 z}{z - 0,75} I_t, \quad (2.18)$$

$PL_i = 59,3$ долл./ед., коэффициент дисконтирования $r = 0,05$, средняя цена микропроцессора в долл линейно связана с объемами их предложения на рынке (в нашем случае с суммарным объемом производства участников в млн ед.):

$$p_t = 178,765 - 0,164 \cdot (Q_{1t} + Q_{2t}) \quad (2.19)$$

(последняя зависимость получена на основе данных о средней цене микропроцессоров и продажах персональных компьютеров на российском рынке). Заданные установившиеся уровни производства участников рынка составляют $Q_{1\infty} = 2$ и $Q_{2\infty} = 10$ млн. микропроцессоров в год. Соответствующие этим уровням значения коэффициентов ρ_i , рассчитанные в соответствии с (2.15), составляют: $\rho_1 = 0,0232$; $\rho_2 = 0,0063$.

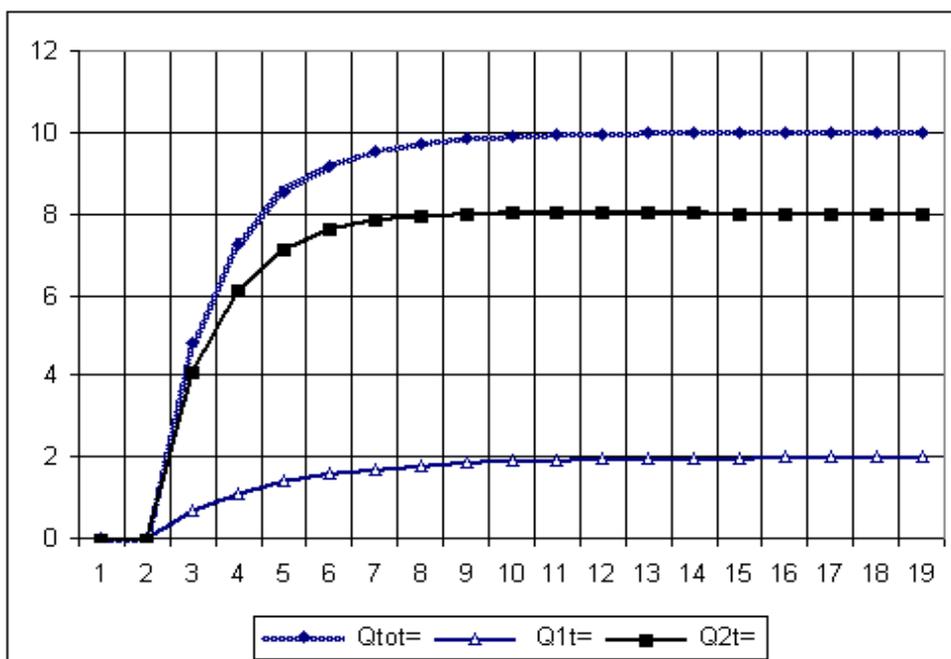


Рис. 2.14. Динамика объемов производства микропроцессоров в случае дуополии

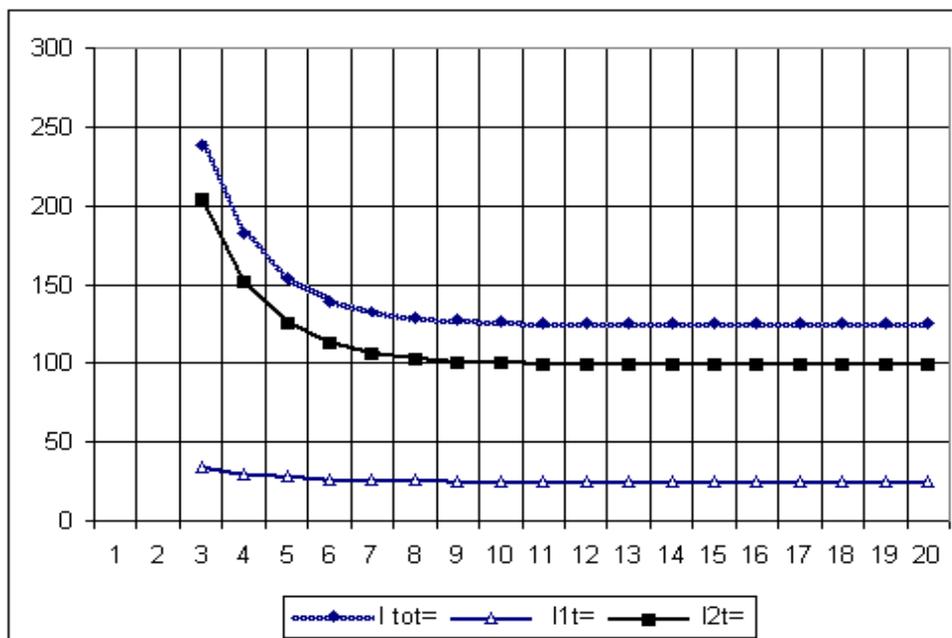


Рис. 2.15. Динамика объемов инвестиций в производства микропроцессоров в случае дуополии

2.7. Заключение

Рассмотренные в главе эконометрические модели позволяют провести содержательный анализ инвестиционных процессов на рынках микроэлектронной продукции. Они создают основу для формирования прогнозных сценариев развития рынков микропроцессоров на среднесрочную перспективу в рамках игрового подхода.

Анализ построенных сценариев свидетельствовал о реальности достижения целей компании ARM Ltd. на рынке микропроцессоров для серверов. В то же время на рынке микропроцессоров x86 для PC, повидимому, не произойдет серьезных структурных сдвигов и изменений объемных показателей.

Разработанные методы и модели могут быть использованы при формировании предложений по созданию структуры отечественного рынка продукции элементной базы НРС-систем, а также по государственной политике стимулирования инвестиционной деятельности в высокотехнологических отраслях.

Список литературы к главе 2

1. Варшавский Л.Е. Исследование динамики показателей рынков наукоемкой продукции (на примере рынков компонентов персональных компьютеров) // Экономика и математические методы. 2004. Т. 40. № 1. С. 101–116.

2. Aizcorbe A., Kortum S. Moores Law and the Semiconductor Industry: A Vintage Model // *Scandinavian Journal of Economics*. 2005. Vol. 107. Iss. 4. Pp. 603–630.
3. Pillai U. A Model of Technological Progress in the Microprocessor Industry // *The Journal Of Industrial Economics*. 2010. Vol. LXI. No. 4. Pp. 877–912.
4. Goettler R.L., Gordon B.R. Does AMD Spur Intel to Innovate More? // *Journal of Political Economy*. 2011. Vol. 119. No. 6. Pp. 1141–1200.
5. Macieira J. Extending the Frontier: A Structural Model of Investment and Technological Competition in the Supercomputer Industry. 2008. URL: ftp://repec.econ.vt.edu/Papers/Macieira/macieira_superc_jan22.pdf
6. Macieira J. A Dynamic Model of Innovation and Technological Competition in the Supercomputer Industry. 2009. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.517.9875&rep=rep1&type=pdf>
7. Варшавский Л.Е. Моделирование динамики ключевых показателей рынков компонентов высокопроизводительных вычислительных систем // *Труды Института системного анализа Российской академии наук*. 2017. Т. 67. № 1. С. 12–27.
8. Варшавский Л.Е. Анализ инвестиционных процессов в ведущих компаниях микроэлектронной промышленности // *Концепции*. 2016. № 1 (35). С. 17–28.
9. Варшавский Л.Е. Экономико-математические методы анализа динамики рыночных процессов и формирования конкурентной среды в микроэлектронной промышленности // *Национальные интересы: приоритеты и безопасность*. 2016. № 336. С. 26–44.
10. Малашевич Б.М., Малашевич Д.Б. Микроэлектроника. Краткие основы и история развития // *Базовые лекции по электронике. Т. II. Твердотельная электроника / под общ. ред. В.М. Пролейко*. М.: Техносфера, 2009. С. 96–170.
11. Kvamme E.F. Life in Silicon Valley: A first-hand view of the region growth // *The Silicon Valley Edge: a habitat for innovation and entrepreneurship / ed. by Chong-Moon Lee et al*. Stanford, Stanford University Press, 2000. Pp. 59–80.
12. Lecuyer C. Fairchild Semiconductor and its influence // *The Silicon Valley Edge: a habitat for innovation and entrepreneurship / ed. by Chong-Moon Lee et al*. Stanford, Stanford University Press, 2000. Pp.158–183.
13. Гроув А. Выживают только параноики. М.: Альпина Бизнес Букс, 2004. 200 с.
14. Moore G. Progress in Digital Integrated Electronics. URL: http://download.intel.com/museum/Moores_Law/Articles-Press_Releases/Gordon_Moore_1975_Speech.pdf
15. Dennard R.H., Gaensslen F.H., Yu H.-N., Rideout V.I., Bassous E., LeBlanc A.R. Design of ion-implanted MOSFET's with very small physical dimensions // *IEEE Journal of Solid-State Circuits*. 1974. SC-9. Pp. 256–268.
16. Варшавский Л.Е. Проблемы повышения энергоэффективности аппаратных средств в области информационных технологий // *Труды ИСА РАН*. 2013. Т. 63. Вып. 3. С. 3–19.
17. Митропольский Ю.И. Элементная база и архитектура будущих суперкомпьютеров // *Микроэлектроника*. 2015. Т. 44. № 3. С. 163–179.
18. Щука А.А. Нанoeлектроника. М.: Физматкнига, 2007.
19. Рудометов Е. Материнские платы и чипсеты. 4-е изд. *Анатомия ПК*. СПб.: Питер, 2007. 368 с.
20. URL: <https://semiengineering.com/uncertainty-grows-for-5nm-3nm/>
21. URL: <https://spectrum.ieee.org/semiconductors/devices/transistors-could-stop-shrinking-in-2021>
22. URL: <https://spectrum.ieee.org/computing/hardware/expect-deeper-and-cheaper-machine-learning>
23. URL: <https://spectrum.ieee.org/tech-talk/computing/hardware/lowbudget-chip-design-how-hard-is-it>

24. URL: <https://spectrum.ieee.org/view-from-the-valley/computing/hardware/move-over-moores-law-make-way-for-huangs-law>
25. URL: <https://www.datanami.com/2018/01/04/2018-will-data/>
26. URL: <https://www.top500.org/news/doe-shifts-exascale-plans-into-high-gear-asks-supercomputing-vendors-to-submit-proposals/>
27. Программа «Цифровая экономика Российской Федерации». Утверждена распоряжением Правительства Российской Федерации от 28 июля 2017 г. № 1632-р.
28. Варшавский Л.Е. Приближенные методы исследования динамики показателей рыночной структуры // Компьютерные исследования и моделирование. 2012. Т. 4. № 1. С. 219–229.
29. Варшавский Л.Е. Использование методов теории управления для формирования рыночных структур // Компьютерные исследования и моделирование. 2014. Т. 6. № 5. С. 839–859.
30. Gordon S. Costs of Adjustment, the Aggregation problem and Investment // *The Review of Economics and Statistics*. 1992. Vol. 74. No. 3. P. 422–429.
31. Basar T., Olsder G.J. *Dynamic Noncooperative Game Theory*. London; New York: Academic Press, 1995.
32. Dockner E.J., Jorgenson S. et al. *Differential Games in Economics and Management Science*. Cambridge: Cambridge University Press, 2000.
33. Кузин Л.Т. Расчет и проектирование дискретных систем управления. М.: Машгиз, 1962.
34. Jury E.I. *Theory and Applications of the Z-Transform Method*. John Wiley, NY, 1964.
35. Варшавский Л.Е. Методы расчета оптимальных по Нэшу стратегий участников динамических игр на основе электронных таблиц // Анализ и моделирование экономических и социальных процессов / Математика. Компьютер. Образование: Сб. научных трудов. Вып. 22. № 3. М.; Ижевск: Институт компьютерных исследований, 2015. С. 84–96.
36. Джексон Т. INTEL. Взгляд изнутри. М.: Лори, 1998.
37. Варшавский Л.Е. Методологические основы моделирования инвестиционного поведения промышленных фирм // Теория и практика институциональных преобразований в России / Ерзнкян Б.Г. (ред.). Вып. 3. М.: ЦЭМИ РАН, 2004. С. 70–96.
38. Варшавский Л.Е. Сравнительный анализ оптимальных по Нэшу–Курно стратегий олигополистов на основе компьютерного моделирования // Теория и практика институциональных преобразований в России / Ерзнкян Б.Г. (ред.). Вып. 19. М.: ЦЭМИ РАН, 2011. С. 11–30.

ГЛАВА 3.

ИССЛЕДОВАНИЕ ПЕРСПЕКТИВНОЙ ДИНАМИКИ ПОКАЗАТЕЛЕЙ РАЗВИТИЯ РЫНКА МОБИЛЬНОЙ СВЯЗИ

3.1. Введение

К настоящему времени мобильная связь проникла во все сферы жизни современного общества. С начала 1990-х гг., когда появились разновидности цифровой технологии мобильной связи 2G и до настоящего времени происходит неуклонное увеличение числа пользователей мобильной связи.

Особенностью развития рынка мобильной беспроводной связи в последнее десятилетие как в США, так и в других странах, является преимущественный рост его масштабов за счет быстрого роста трафика данных по мере освоения технологии 3G и особенно перехода на технологию 4G. Так, в 2012 г. на долю технологии 4G, осваивающейся с 2010 г., приходилось 14% всего мобильного трафика данных в мире, хотя ее доля в общем числе подключений (connections) составляла всего 1% [2]. В 2016 г. удельный вес этой технологии в общемировом мобильном трафике данных составляла уже 69% при доле подключений 26% [3].

Переход к технологии 4G позволил существенно повысить объем мобильного трафика данных. Так, в 2017 г. объем мобильного трафика данных в России превысил, по нашим оценкам, в 6 раз уровень этого показателя в 2011 г. В США объем мобильного трафика данных составил 15 687 млрд МВ (рис. 3.1), что в 40 раз выше, чем в 2010 г. Если в 2011 г. доля мобильного трафика в общих доходах компаний-провайдеров США составляла немногим более одной трети, то уже в 2014 г. она превысила 50% [4]. Интересно отметить, что в США в последние годы соотношение между общим числом подключений Q_{it} (connections, subscribers) и десятичным логарифмом трафика $\log_{10}(Traffic_{it})$ остается относительно стабильным, и в среднем в 2013–2017 гг. составляло $Q_t / \log_{10}(Traffic_t) = 99,5$ млн соединений на $\log(MB)$.

По оценкам в ближайшие несколько лет доля мобильного трафика данных в общих доходах американских компаний-провайдеров повысится до 75% [5, 6]. Дальнейший рост этой доли возможен при переходе к технологии 5G, хотя такой переход потребует решения ряда серьезных проблем.

Рост активности в отрасли мобильной связи в разных странах мира сопровождается повышением рыночной доли крупных компаний-провайдеров. Так, по данным компании Мегафон, в нашей стране на долю 4 провайдеров (Мегафон, МТС,

Билайн, Теле2) приходится почти 99% обслуживаемых в стране абонентов мобильной связи. В США на долю 4-х компаний (Verizon, AT&T, T-Mobile и Sprint) также приходится около 100% всех подключений и доходов от мобильного сервиса в США, причем доля этих компаний в подключениях в сложившейся квадрополии составляла в 2017 г. соответственно 35,5%, 34,3, 17,3 и 12,8%.

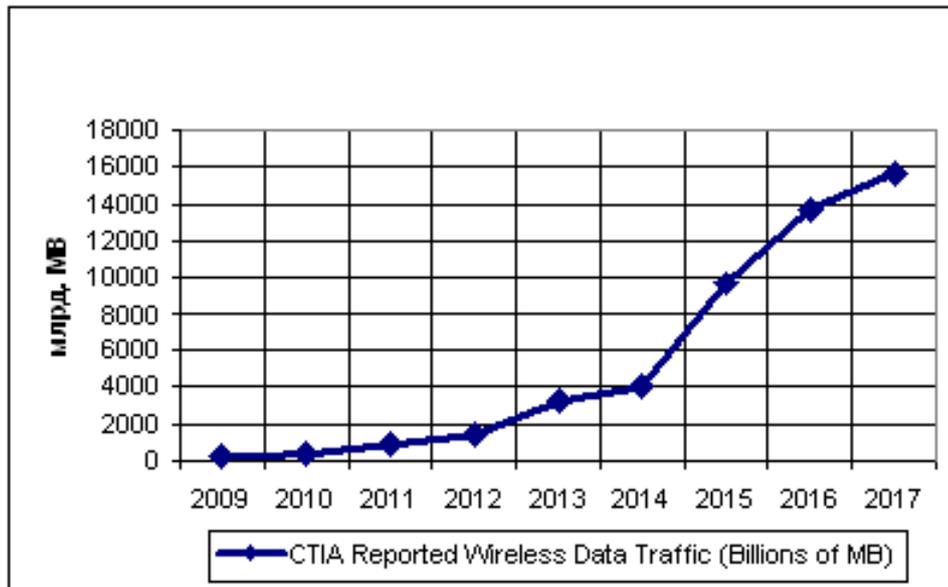


Рис. 3.1. Мобильный трафик данных в США в млрд МВ (составлено по данным СТИА)

В конце 2010 г. в Федеральной комиссии США по связи (*Federal Communications Commission*, сокр. FCC) обсуждалась проблема слияния T-Mobile и Sprint! Противники объединения опасаются усиления концентрации на рынке всего 3-х компаний (возникнет триполия). В свою очередь, сторонники слияния отмечали в качестве положительного факта то, что объединенная компания будет более сильным конкурентом для AT&T и Verizon. Кроме того, они считали, что триполия будет способствовать более быстрому развертыванию системы 5G (см., напр. [7]).

Проблема формирования эффективной структуры рынка мобильных сервисов стоит и в нашей стране, что, в частности, связано с подготовкой к переходу к капиталоемкой технологии 5G. В связи с этим, представляет практический интерес разработка методических подходов к исследованию перспективной структуры этого рынка.

В настоящей главе на основе динамической игровой модели проводится исследование перспективной динамики показателей рынка мобильной связи (на примере США, в условиях как существующей квадрополии ($N = 4$), так и возможной триполии ($N = 3$)). Материал данной главы основан на статье автора [1].

3.2. Структура модели

Модель состоит из следующих блоков:

1) динамики соотношения между объемами капитальных вложений (*Capex* в млн долл.) и числом мобильных подключений Q_{it} в i -й компании (в млн):

$$Q_{it} = W_i(z) u_{it} + Q_{0it} = \frac{\alpha_i z}{(z - \lambda_i)^2} u_{it} + Q_{0it}, \quad (3.1)$$

где z – оператор сдвига, т.е. $zx_t = x_{t+1}$, $u_{it} = \text{Capex}_{it}$, Q_{0it} – число мобильных подключений, не связанное с капитальными вложениями, $0 < \lambda_i < 1$, $i = 1, 2, \dots, N$.

2) обратной функции спроса, связывающей среднегодовую цену в олигополии P_t (в долл.) с суммарным числом подключений (в млн):

$$P_t = a - b \sum_i Q_{it}. \quad (3.2)$$

Предполагается, что олигополисты используют скользящее планирование и в каждый момент времени τ максимизируют чистую текущую стоимость (*NPV*):

$$J_{\tau i} = \sum_{t=\tau}^{\tau+Tp} \beta^t \left[(f_i m_i p_t) Q_{it} - u_{it} - \frac{1}{2} \rho_i u_{it}^2 \right] \rightarrow \max_{u_{it}}, \quad (3.3)$$

где β – дисконт-фактор; f_i – соотношение между средними доходами на одно соединение в i -й компании и соответствующим показателем для суммы компаний, m_i – среднее значение EBITDA margins (средняя доля прибыли в доходах без учета процентных платежей, налогов, амортизации; англ. earnings before interest, tax, depreciation and amortization) в i -й компании; ρ_i – коэффициенты при затратах регулирования (adjustment costs), характеризующих наличие инвестиционных ограничений. Управляющими переменными в модели (3.1)–(3.3) являются капитальные вложения u_t . Задача состоит в определении оптимальных по Нэшу игровых стратегий компаний-провайдеров.

Следует отметить, что идентификация составных блоков рассматриваемой модели рынка мобильной связи в США (как, впрочем, и в других странах) сопряжена с трудностями, обусловленными различием статистических данных, приводимых как официальными, так и частными организациями. В настоящей работе использовались данные годовых докладов Федеральной комиссии по связи США за 2009–2017 гг. Оценки параметров соотношений (3.1) и (3.2), полученные путем использования соответственно нелинейного и линейного метода наименьших квадратов, приведены в табл. 3.1 и 3.2 (в нижней строке табл. 3.2 указаны значения среднеквадратических отклонений оценок параметров). В качестве оценок m_i приняты

среднегодовые значения удельной прибыли за 2013–2017 гг. (табл. 3.3). Для получения оценок коэффициентов ρ_i использован подход, изложенный в главе 2 (см. соотношение (2.15), а также в статьях [8, 9]).

Таблица 3.1

Значения параметров зависимости (3.1)

| Компания | α_i | λ_i | Коэффициент вариации v |
|------------------|------------|-------------|--------------------------|
| Verizon | 1,198 | 0,739 | 0,085 |
| AT&T | 1,118 | 0,739 | 0,085 |
| T-Mobile | 4,422 | 0,502 | 0,064 |
| Sprint | 0,654 | 0,752 | 0,024 |
| T-Mobile +Sprint | 1,666 | 0,708 | 0,045 |

Таблица 3.2

Значения параметров зависимости (3.2)

| a | b | R^2 |
|--------|---------|-------|
| -1,012 | 856,110 | 0,939 |
| 0,097 | 33,314 | |

Таблица 3.3

Принятые в модели значения m_i (EBITDA Margins)

| Verizon | AT&T | T-Mobile | Sprint |
|---------|-------|----------|--------|
| 0,530 | 0,447 | 0,287 | 0,266 |

Возможность использования эконометрических моделей при прогнозировании показателей быстро развивающегося исследуемого рынка обусловлена тем, что значительную часть инфраструктуры, созданной в США для технологии связи 4G, предполагается использовать и для технологии 5G, на распространение которой в ближайшее время ориентированы компании и-провайдеры [10].

Ниже на основе игровой модели (3.1)–(3.3) исследуется динамика показателей олигополистического рынка мобильной связи США на дальнюю перспективу (2030 г.), соответствующая следующим сценариям: 1) обеспечение компаниями-провайдерами наиболее выгодной по критерию NPV траектории достижения гипотетических целевых уровней подключений к 2030 г. $Q_{i\infty}^*$, которые определяются, как

$$Q_{i\infty}^* = \frac{\gamma_i}{(1 - \lambda_i)^2} u_{iavg}, \text{ где } u_{iavg} - \text{ среднегодовые за последние 7–8 лет объемы капитальных}$$

вложений (экстраполяционный сценарий); 2) достижение желаемых гипотетических рыночных долей компаний к 2030 г., согласованных с их инвестиционными возможностями и оптимизацией NPV , при различных сценариях роста суммарного

по стране мобильного трафика данных (является экзогенной переменной). Во всех сценариях в соответствии с соотношением (2.15) главы 2 подбираются коэффициенты ρ_i .

В результате расчетов на основе расширенных уравнений Риккати для линейных динамических игр с квадратичным критерием [11] определялись разомкнутые (open-loop) оптимальные по Нэшу стратегии участников рынка.

3.3. Прогнозные сценарии

а) Сохранение квадрополии

В экстраполяционном сценарии оптимальная в соответствии с критерием *NPV* динамика числа подключений 4-х компаний-провайдеров, согласованная с заданными на 2030 г. целями (соответственно 174 млн, 166 млн, 70 млн и 43 млн подключений), представлена на рис. 3.2. При этих гипотетических целях компаний рост числа подключений будет иметь место лишь в компаниях-лидерах (Verizon, AT&T). Ввиду невысоких объемов капитальных вложений в компании Sprint, число подключений в этом провайдере должно уменьшиться с 59 млн в 2017 г. до 43 млн в 2025 г. В компании T-Mobile в соответствии с принятой гипотезой, уровень подключений, несмотря на рост капитальных вложений в последние годы, должен немного снизиться (до 70 млн).

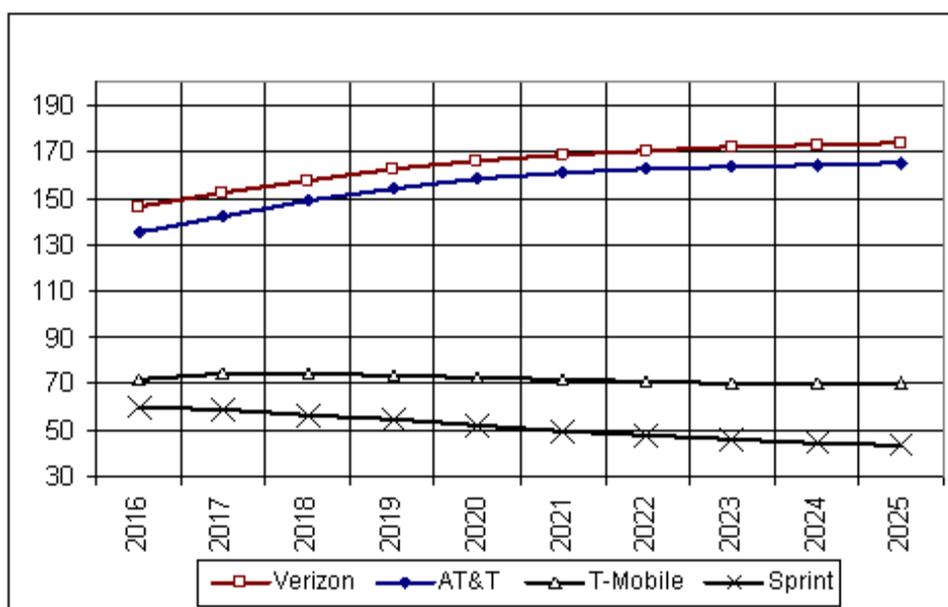


Рис. 3.2. Динамика числа подключений 4-х компаний-провайдеров, согласованная с гипотетическими целями на 2030 г. (экстраполяционный сценарий)

Оптимальная динамика показателей компаний при различных сценариях роста мобильного трафика данных (с темпами прироста $\gamma = 5\%$, 10% , 15%),

соответствующая сохранению рыночных долей компаний по показателю подключений к 2030 г. на уровне 2017 г. (соответственно 0,355; 0,343; 0,173; 0,128), представлена на рис. 3.3, 3.4. При этом использовано следующее соотношение между суммарными в четырех компаниях числом подключений в млн и мобильным трафиком (в млрд МВ): $Q_{\Sigma t} = 99,5 \cdot \log_{10}(Traffic_{\Sigma t})$.

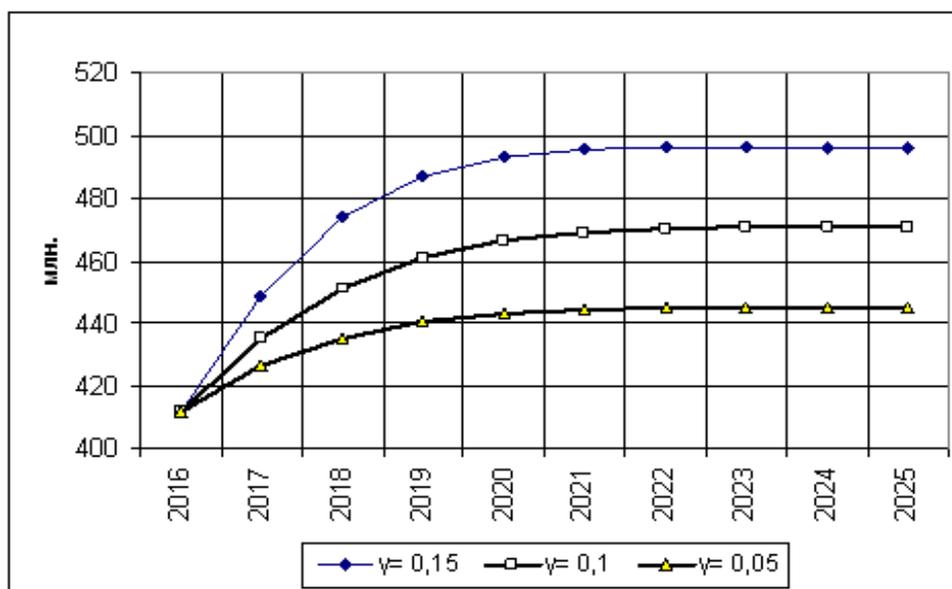


Рис. 3.3. Динамика суммарного числа подключений 4-х компаний-провайдеров при различных сценариях (вариантах) роста трафика данных

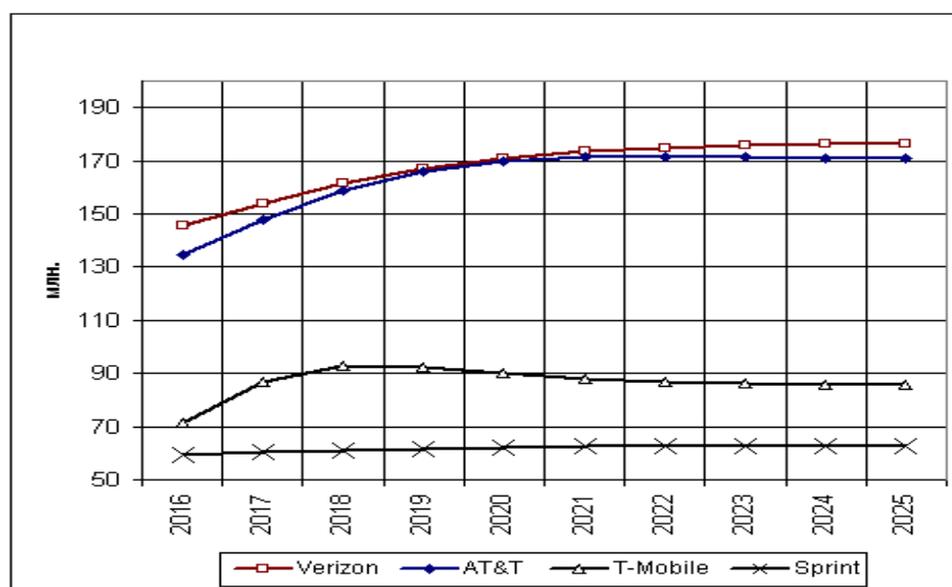


Рис. 3.4. Динамика числа подключений 4-х компаний-провайдеров при сценарии 15%-го роста трафика данных

Следует отметить близость траекторий показателя подключений как для компаний Verizon и AT&T в сценарии 15%-го роста мобильного трафика и в экстраполяционном сценарии.

В связи с меньшим числом суммарных по совокупности 4 компаний подключений в экстраполяционном сценарии по сравнению со сценарием 15%-го роста мобильного трафика, удельные доходы на 1 подключение ARPU (average return per unit) в первом сценарии существенно выше, чем во втором. Интересно отметить, что критерий максимума NPV для компаний приводит в первые годы к меньшему значению ARPU, чем критерий максимума доходов Rev (рис. 3.5).

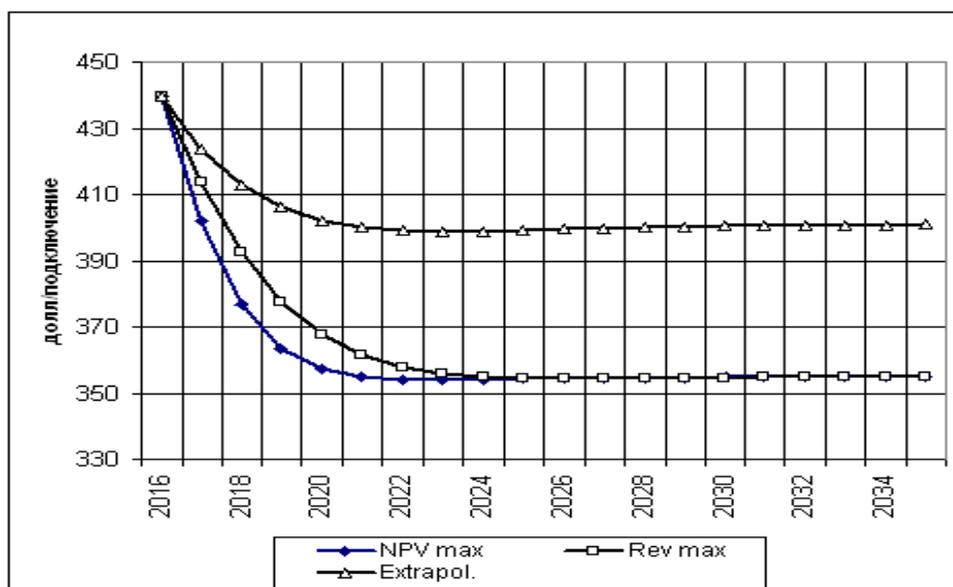


Рис. 3.5. Динамика ARPU (долл) при различных возможных критериях оптимизации компаний-провайдеров (максимум NPV , максимум доходов Rev max в сценариях 15%-го роста трафика) и при экстраполяционном сценарии

б) Переход к триополии

Реализация исследуемых сценариев в триополии (Verizon, AT&T, T-Mobile + Sprint) может привести к более равномерному распределению числа подключений, чем в случае квадрополии. Это связано с возможностью проявления синергического эффекта при объединении T-Mobile, наращивающей в последние годы объемы капиталовложений, и Sprint (рис. 3.6, 3.7).

Следует однако отметить, что при рассматриваемом относительно умеренном темпе прироста мобильного трафика в 15% и соответствующих ему целевых уровнях подключений компаний (соответственно 176 млн, 170 млн и 149 млн) оптимальная по критерию NPV (3.3) стратегия объединенной компании может быть достигнута лишь при достаточно высокой марже (порядка 40%, т.е. при $m_3 = 0,4$). Это значение маржи для T-Mobile+Sprint принято в дальнейших расчетах.

Интересно также отметить близость динамики и числа подключений в сценарии 15%-го роста мобильного трафика и в экстраполяционном сценарии как для компании Verizon, так и для объединенной компании – Mobile+Sprint (рис. 3.6, 3.7).

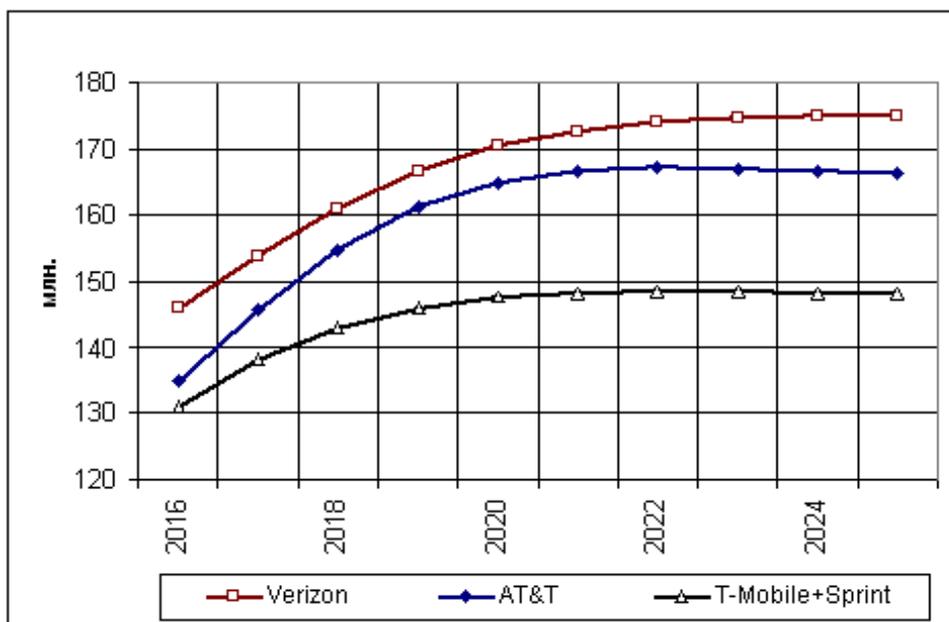


Рис. 3.6. Динамика числа подключений в триополии при экстраполяционном сценарии

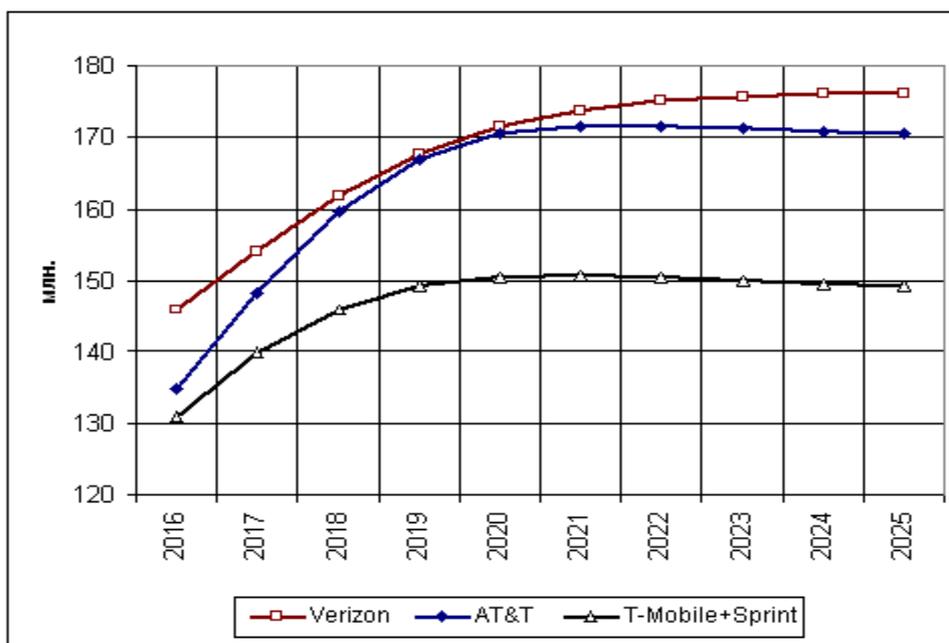


Рис. 3.7. Динамика числа подключений в компаниях триополии при сценарии 15%-го роста трафика данных

Вместе с тем, расчеты показывают, что новая объединенная компания могла бы захватить наибольшую долю рынка, отклонившись от целей, ожидаемых другими компаниями. Так, при ориентации этой компании на достижение 35%-й доли на рынке в 2030 г. (вместо предполагаемой другими участниками рынка 30,1%, при целевых установках конкурентов на прежних уровнях: 35,5 и 34,3%) ее рыночная доля уже к 2025 г. составила бы более 36% (рис. 3.8, вариант 1).

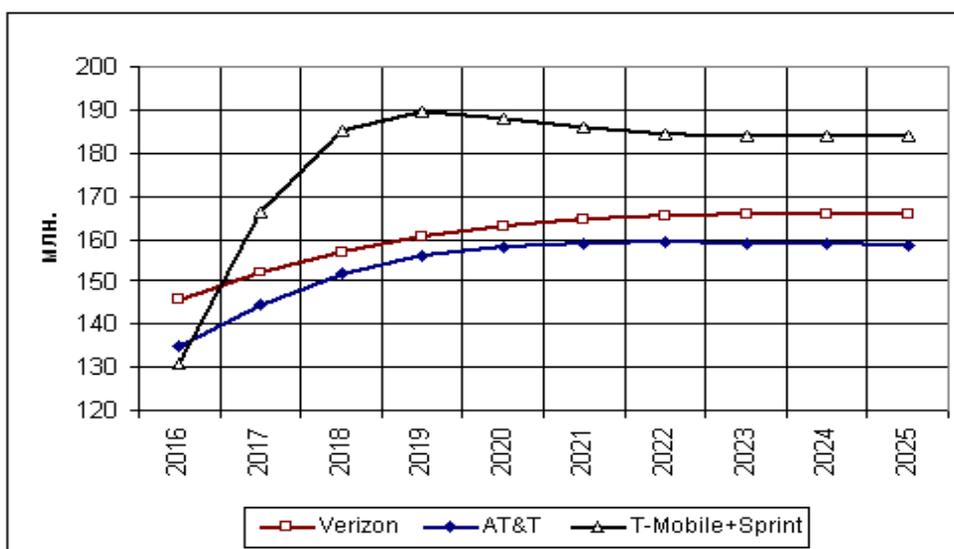


Рис. 3.8. Динамика числа подключений в компаниях триополии при отклонении целей компаний от первоначально принятых (вариант 1, сценарий 15%-го роста трафика данных)

При другом варианте событий, когда Verizon попытается, достигнуть 40% доли рынка, а AT&T и T-Mobile+Sprint – 35%-й доли рынка каждый (полагая, что конкуренты будут ориентироваться на свои старые цели), новая объединенная компания, увеличив свою рыночную долю до 33,2%, несколько потеснит AT&T. Рыночная доля последней составит 29,4% (рис. 3.9, вариант 2). Таким образом, превращение квадрополии мобильной связи в США в триополию может существенно повлиять на структуру рынка в сторону повышения рыночной доли объединенной компании.

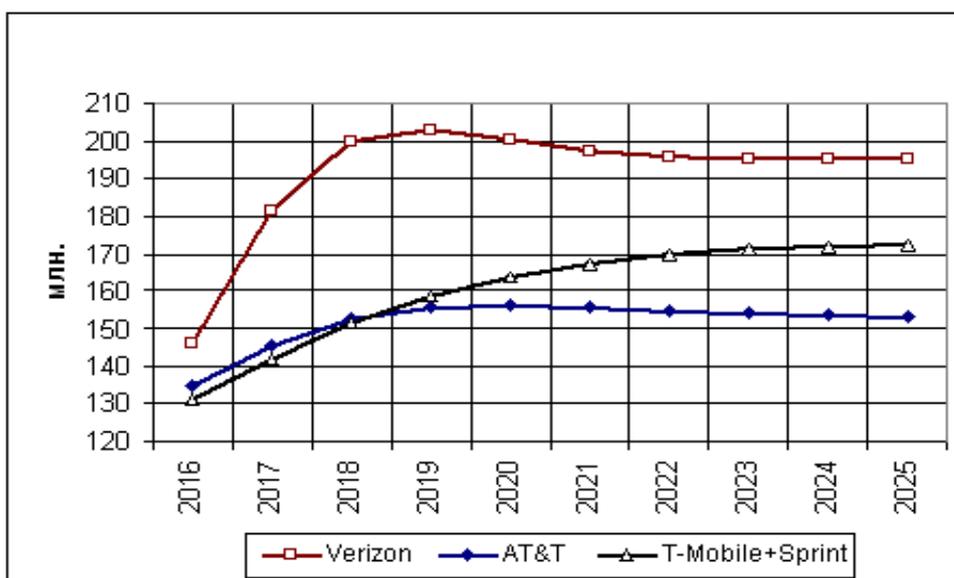


Рис. 3.9. Динамика числа подключений в компаниях триополии при отклонении целей компаний от первоначально принятых (вариант 2, сценарий 15%-го роста трафика данных)

Иными словами, опасения компаний-лидеров появлением нового мощного соперника не лишены оснований! Вместе с тем, многое зависит от того, способна ли будет новая объединенная компания улучшить показатели своей эффективности, в особенности удельную прибыль m_3 (EBITDA margin).

3.4. Некоторые препятствия на пути широкомасштабного развертывания новых технологий мобильной связи в США

В последнее время значительные усилия компаний-провайдеров в США и во многих странах мира направлены на развертывание технологии мобильной связи 5G. Необходимость перехода к 5G связывается с увеличением трафика, предполагаемым бурным распространением Интернета вещей IoT, и расширением масштабов обслуживания движущихся объектов и др. При этом, по сравнению с технологией 4G, время задержки при передаче сигналов должно примерно на 2 порядка уменьшиться, а скорость передачи сигналов увеличиться более, чем в 20 раз.

Для передачи сигналов с повышенной спектральной эффективностью (spectrum efficiency) будут использованы миллиметровые волны, которые, однако, поглощаются каплями дождя и частицами в воздухе, и не могут проходить через преграды. Поэтому предполагается примерно через каждые 250–500 м сооружать небольшие базовые станции (micro basestations), что резко увеличит плотность их установки (в десятки и сотни раз). Так, по расчетам компании UBS в Нью-Йорке придется увеличить число базовых станций в 500 раз (с 1251 до 635 639)!

Однако развертывание системы мобильной связи 5G требует тщательной проработки как с позиций экономической эффективности (особенно спроса), так и с позиций влияния значительного увеличения плотности и интенсивности сверхвысокочастотного электромагнитного излучения на здоровье населения и на инфраструктуру (подробнее см. главу 7). В связи с вышеприведенными обстоятельствами, в перспективе освоение технологии 5G, а затем и 6G может столкнуться с существенными препятствиями.

3.5. Заключение

1. Использование линейных динамических игр с квадратичным критерием позволяет провести содержательный анализ стратегий компаний-провайдеров и разработать сценарии и прогнозы динамики показателей олигополистических рынков мобильной связи как в России, так и в зарубежных странах.

2. Развертывание системы мобильной связи 5G требует тщательной проработки как с позиций экономической эффективности (особенно спроса), так и с позиций влияния значительного увеличения плотности и интенсивности сверхвысокочастотного электромагнитного излучения на здоровье населения и на инфраструктуру.

Список литературы к главе 3

1. Варшавский Л.Е. Исследование показателей развития рынка мобильной связи // Концепции. 2019. № 1 (38). С. 36–45.
2. Maternia M. Deployment and Transition of Telecommunication Systems // In the Telecommunications Handbook: Engineering Guidelines for Fixed, Mobile and Satellite Systems, Edited by J.T.J. Penttinen. John Wiley & Sons, 2015.
3. URL: <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html>
4. URL: <https://www.ctia.org/the-wireless-industry/infographics-library>
5. Annual Report and Analysis of Competitive Market Conditions with Respect to Mobile Wireless, Including Commercial Mobile Services. Sixteenth Report. Federal Communications Commission FCC 13-34. Released: March 21, 2013.
6. URL: <https://www.reuters.com/article/us-usa-mobilephone-data/u-s-mobile-data-revenue-surpasses-voice-call-revenue-analyst-idUSBREA2C24J20140313>
7. T-Mobile/Sprint Merger Offers Public Interest Benefits: Likely Presents a Fast Track to 5G // Perspectives from FSF Scholars. 2018. July 12. Vol. 13. No. 27. URL: http://www.freestatefoundation.org/images/T-Mobile-Sprint_Merger_Offers_Public_Interest_Benefits_071218.pdf
8. Варшавский Л.Е. Экономико-математические методы анализа динамики рыночных процессов и формирования конкурентной среды в микроэлектронной промышленности // Национальные интересы: приоритеты и безопасность. 2016. № 336. С. 26–44.
9. Варшавский Л.Е. Моделирование динамики ключевых показателей рынков компонентов высокопроизводительных вычислительных систем // Труды Института системного анализа Российской академии наук. 2017. Т. 66. № 1. С. 12–27.
10. 5G: The new network arrives TMT Predictions 2019. URL: <https://www2.deloitte.com/insights/us/en/industry/technology/technology-media-and-telecom-predictions/5g-wireless-technology-market.html>.
11. Basar T., Olsder G.J. Dynamic Noncooperative Game Theory. London; New York: Academic Press, 1995.

РАЗДЕЛ 2.

ТЕХНИКО-ЭКОНОМИЧЕСКИЙ АНАЛИЗ ТЕНДЕНЦИЙ РАЗВИТИЯ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

ГЛАВА 4. ИССЛЕДОВАНИЕ ДИНАМИКИ КЛЮЧЕВЫХ ПОКАЗАТЕЛЕЙ ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ СУПЕРКОМПЬЮТЕРОВ

4.1. Введение

В настоящее время использование высокопроизводительных вычислений (high-performance computing, HPC) является неременным условием научно-технического развития. Важным сегментом рынка высокопроизводительных вычислений является быстро развивающийся рынок суперкомпьютеров, объем которого уже в 2016 г. превысил 4 млрд долл. [1].

Высокий спрос на суперкомпьютеры во многом обусловлен стремительным увеличением их производительности. Так, если в 1997 г. был превышен рубеж производительности суперкомпьютеров в 1 терафлопс, то в 2008 г. – 1,10 петафлопс (Пфлопс). Согласно рейтингу Top 500, с июня 2016 г. до июня 2017 г. наиболее производительным в мире являлся китайский суперкомпьютер Sunway TaihuLight, на котором достигнута скорость выполнения операций 93 Пфлопс. Наконец, в мае текущего 2022 г. самым высокопроизводительным суперкомпьютером из списка Top 500 стал суперкомпьютер Frontier (США) с производительностью 1,102 экзафлопс.

Самые производительные суперкомпьютеры сконцентрированы в исследовательских центрах государственных структур и крупных компаний. Так, 59 из 117 суперкомпьютеров производительностью свыше 1 Петафлопс (Пфлопс)^{vii}, входивших в ноябре 2016 г. список 500 наиболее производительных суперкомпьютеров мира Top 500 (он публикуется 2 раза в год [2]), эксплуатировалось в исследовательских центрах.

Однако суперкомпьютеры все более широко используются не только в исследовательском и академическом секторах, но и в бизнесе. Так, если в 1997 г. в Top 500, входил 161 суперкомпьютер из коммерческого сектора (industry), то в 2007 г. –

^{vii} В списке Top 500 приводится ранжировка суперкомпьютеров по производительности R_{\max} в соответствии с тестом Linpack, связанным с решением линейных уравнений большой размерности.

уже 248. Примерно такое же число суперкомпьютеров из этого сектора было представлено и в ноябрьском списке Top 500 за 2016 г. (245 суперкомпьютеров). Драйверами активного применения суперкомпьютеров в бизнесе являются: 1) рост затрат на создание и испытание образцов (прототипов) изделий; 2) резкое снижение затрат на виртуальное моделирование и вычисления [3].

Следует отметить, что в последние 2 десятилетия «...рост производительности суперкомпьютеров происходит в равной мере, как за счет повышения быстродействия физических элементов, так и за счет увеличения числа элементов в вычислительной установке; последнее связано с развитием многоуровневого параллелизма обработки» [4, с. 578]. В связи с этим, актуальными являются исследование развития как физико-технологической базы суперкомпьютеров, основу которой составляют устройства микроэлектроники, так и их архитектуры.

В настоящей главе проводится экономико-математический анализ тенденций изменения ключевых показателей высокопроизводительных суперкомпьютеров, динамики их распространения и изменения структуры, а также факторов, влияющих на их энергоэффективность. Исследуются подходы к первоначальной оценке возможностей создания суперкомпьютера экзафлопсной производительности. Материал главы основан на статьях автора [5, 6].

4.2. Анализ ключевых технико-экономических показателей оборудования для суперкомпьютеров

4.2.1. Динамика распространения высокопроизводительных суперкомпьютеров

Последние 3 десятилетия ознаменовались ускоренным технологическим развитием микроэлектроники и средств вычислительной техники. Существенно улучшились все основные параметры, характеризующие качественные и количественные характеристики микроэлектронных схем, долговременной памяти и создаваемых на их основе вычислительных устройств.

Развитие физико-технологической базы существенно повлияло на архитектуру и характеристики суперкомпьютеров. Так, в 1980–1990 гг. при построении суперкомпьютеров использовалась векторно-конвейерная структура, в рамках которой разные процессоры объединялись с общей памятью. Производительность систем составляла тогда 1–10 Гигафлопс. В 1990–1995 гг. за счет применения массово-параллельных структур с распределенной памятью (MPP-MIMD), основанных на КМОП технологии, удалось выйти на уровень производительности 100 Гигафлопс и более.

С середины 1990-х гг. построение суперкомпьютеров осуществлялось «...путем объединения крупносерийных микропроцессоров широкого применения» [4]. В 1997 г. был превышен рубеж производительности суперкомпьютеров в 1 терафлопс, в 2005 г. – 280 терафлопс, а в 2008 г. – 1,10 петафлопс (Пфлопс). В 2016 г. суммарная по показателю R_{\max} производительность 500 суперкомпьютеров из списка Top 500 составила 672,1 Пфлопс. В мае текущего 2022 г. суперкомпьютером Frontier (США) была превышена производительность в 1 экзафлопс, а суммарная производительность 500 суперкомпьютеров из списка Top 500 превысила 4.4 экзафлопс (<https://www.hpcwire.com/2022/05/30/the-final-frontier-us-has-its-first-exascale-supercomputer>).

Вместе с тем, необходимо отметить, что скорость освоения все более высокой производительности по показателю R_{\max} суперкомпьютерами из списка Top 500 уменьшается. Так, уровень производительности в 1 терафлопс был достигнут всеми 500 суперкомпьютерами из этого списка в 2005 г., производительности в 100 терафлопс – в 2013 г., т.е. также за 8 лет. Однако за такой же период времени, т.е. к концу 2016 г., уровень производительности в 1 Пфлопс был превышен лишь на 117 суперкомпьютерах. Лишь в 2019 г. производительность всех суперкомпьютеров из списка Top 500 превысила 1 Пфлопс.

О динамике освоения высоких уровней производительности дает представление рис. 4.1, на котором приведена динамика доли f_t вычислительных мощностей суперкомпьютеров производительностью более 1 терафлопс, 100 терафлопс и 1 петафлопс в суммарной вычислительной мощности 500 суперкомпьютеров (в соответствии с показателем R_{\max}).

Для количественной оценки освоения высоких уровней производительности суперкомпьютеров в настоящей работе использовалась логистическая зависимость,

$$\frac{df}{dt_t} = \alpha f(1 - f), \quad (4.1)$$

которая может быть приведена к следующему виду [7]:

$$\ln(f_t / (1 - f_t)) = \alpha t + \beta, \quad (4.2)$$

где α характеризует скорость освоения уровней производительности суперкомпьютеров. Сопоставление оценок α этой зависимости для разных уровней производительности R_{\max} свидетельствует о замедлении освоения достигаемых рекордных уровней (табл. 4.1; в скобках под значениями оценок здесь и далее приведены значения среднеквадратических ошибок оценок коэффициентов). Данная тенденция свидетельствовала об отсутствии спроса или о неготовности широкого круга

потребителей к эксплуатации высокопроизводительных компьютеров петафлопсного уровня в 2015–2017 гг.

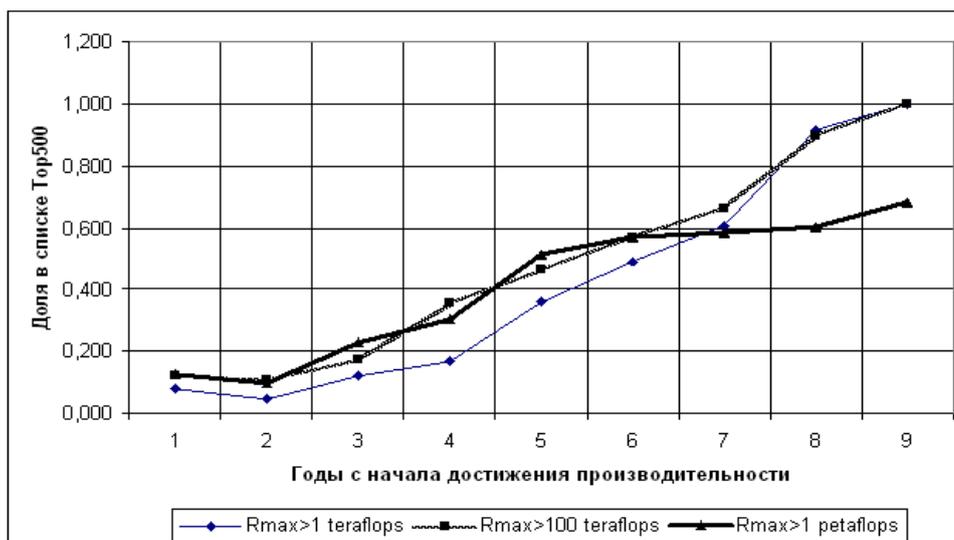


Рис. 4.1. Динамика освоения уровней производительности в 1 терафлопс, 100 терафлопс и 1 петафлопс суперкомпьютерами из списка Top 500

Таблица 4.1

Оценки коэффициентов зависимости (4.1)

| | α | β | R^2 | F |
|----------------------------|------------------|-------------------|-------|---------|
| $R_{\max} > 1$ teraflops | 0,692 (0,088) | -3,975 (0,447) | 0,911 | 61,144 |
| $R_{\max} > 100$ teraflops | 0,581 (0,057) | -3,027 (0,287) | 0,946 | 104,939 |
| $R_{\max} > 1$ petaflops | 0,380 (0,047) | -2,378 (0,361) | 0,905 | 66,641 |

4.2.2. Динамика производительности суперкомпьютеров

В настоящем пункте представлены результаты проведенного автором экономико-статистического анализа факторов, влияющих на производительность мощных суперкомпьютеров (применительно к суперкомпьютерам петафлопсного класса $R_{\max} > 1$ Пфлопс). В табл. 4.2 представлены коэффициенты корреляции производительности суперкомпьютеров и их основных технических характеристик ($R_{\max} > 1$ Пфлопс) с 2010 по 2016 г. Полученные данные подтверждают очевидное положительное влияние на производительность таких факторов, как электрическая мощность, количество ядер процессоров, а также ядер-ускорителей (сопроцессоров). Также обращает на себя внимание слабая корреляция показателей производительности и внутренней частоты процессоров. Такая ситуация во многом связана с различными подходами к архитектуре и составу оборудования, использованными в суперкомпьютерах, которые обусловлены спецификой решаемых задач.

Таблица 4.2

Динамика коэффициентов корреляции производительности суперкомпьютеров и их основных технических характеристик ($R_{\max} > 1$ Пфлопс)

| Годы | Производительность R_{\max} | Электрическая мощность | Количество ядер процессоров | Количество ядер-ускорителей (со-процессоров) | Внутренняя частота процессоров |
|------|-------------------------------|------------------------|-----------------------------|--|--------------------------------|
| 2010 | 1,000 | 0,425 | 0,616 | 0,753 | 0,284 |
| 2011 | 1,000 | 0,899 | 0,984 | -0,191 | -0,482 |
| 2012 | 1,000 | 0,746 | 0,863 | 0,424 | -0,348 |
| 2013 | 1,000 | 0,870 | 0,948 | 0,809 | -0,255 |
| 2014 | 1,000 | 0,983 | 0,949 | 0,827 | -0,188 |
| 2015 | 1,000 | 0,886 | 0,943 | 0,803 | -0,239 |
| 2016 | 1,000 | 0,537 | 0,971 | 0,266 | -0,312 |

По данным ноябрьского выпуска Top 500 за 2016 г. для выборки суперкомпьютеров производительностью свыше $R_{\max} > 1$ Пфлопс построена следующая регрессионная зависимость, связывающая производительность (в Гигафлопс) с числом ядер PR_i и электрической мощностью W_e (в кВт):

$$R_{\max} = 8,454 PR + 184,243 W_e + 1178011,376, \quad R^2 = 0,947. \quad (4.3)$$

(0,232)
(70,904)
(271092,052)

Расчеты, проведенные по данным за 2005–2016 гг. показывают более сильное влияние числа ядер на производительность суперкомпьютеров петафлопсного класса, чем на производительность менее мощных машин (так, эластичность этой группы компьютеров $a = 1,110$ почти в полтора раза превосходит эластичность группы суперкомпьютеров производительностью $R_{\max} > 100$ терафлопс, равную 0,739).

4.2.3. Динамика структуры суперкомпьютеров по их архитектуре

Наряду с анализом динамики производительности, значительный интерес представляет исследование динамики структуры суперкомпьютеров по их архитектуре. В настоящее время в списке Top 500 присутствуют суперкомпьютеры с массивно-параллельной (Massively parallel processing – MPP) и кластерной (Cluster) архитектурой. Суперкомпьютеры с массивно-параллельной архитектурой MPP характеризуются в среднем более высокой производительностью, чем кластерные (до 2013 гг. они регулярно занимали 1-е место в рейтингах Top 500). Однако особенностью последнего десятилетия является уменьшение их доли в списке 500 наиболее производительных вычислительных машин. Так, доля суперкомпьютеров с архитектурой MPP в общем числе суперкомпьютеров со 100% в 2005 г. до 53% в 2010 г. и 13,6% в 2016 г. Доля этих же суперкомпьютеров в суммарной производительности всех суперкомпьютеров снизилась со 100% в 2005 г. до 61% в 2010 г. и 41% в 2016 г.

В настоящей главе исследовалась динамика структуры суперкомпьютеров по их архитектуре. С этой целью по данным за 2008–2016 гг. была построена следующая динамическая модель конкуренции в форме векторной авторегрессии (VAR):

$$\begin{pmatrix} \Delta \ln(R_{\max t}^{Cluster}) \\ \Delta \ln(R_{\max t}^{MPP}) \end{pmatrix} = \begin{pmatrix} \varphi_1 & \theta_1 \\ \varphi_2 & \theta_2 \end{pmatrix} \begin{pmatrix} \ln(R_{\max t-1}^{Cluster}) \\ \ln(R_{\max t-1}^{MPP}) \end{pmatrix} + \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}, \quad (4.4)$$

где $\ln(R_{\max t}^{Cluster})$, $\ln(R_{\max t}^{MPP})$ – натуральные логарифмы суммарной производительности совокупности суперкомпьютеров с архитектурой Cluster и MPP; Δ – знак приращения во времени; φ_i, θ_i, d_i – постоянные коэффициенты. Параметры построенной модели представлены в табл. 4.3.

Таблица 4.3

Оценки параметров φ_i, θ_i, d_i модели (4.4), $i = 1, 2$

| Совокупность суперкомпьютеров с архитектурой | φ_i | θ_i | d_i | Коэффициент детерминации R^2 |
|--|-------------------|------------------|------------------|--------------------------------|
| Cluster | -0,103 (0,037) | 0,000 | 0,894 (0,149) | 0,523 |
| MPP | -1,362 (0,720) | 1,225 (0,522) | 0,421 (3,652) | 0,656 |

Система (4.4) имеет устойчивое положение равновесия, характеризуемое координатами: $\ln(R_{\max t}^{Cluster}) = 8,717$ и $\ln(R_{\max t}^{MPP}) = 8,149$, что соответствует суммарной производительности суперкомпьютеров с архитектурой Cluster и MPP: $R_{\max t}^{Cluster} = 6104,4$ Пфлопс и $R_{\max t}^{MPP} = 3460,3$ Пфлопс. Положению равновесия соответствуют доли архитектур Cluster и MPP в совокупной производительности, равные соответственно 63,8 и 36,2%. В 2016 г. эти доли составляли 58,9 и 41,1%.

Расчеты на основе модели (4.4) показали, что при сохранении существовавших тенденций, в 2020 г. доля мощных суперкомпьютеров архитектуры Cluster должна была составить 62,6%, а архитектуры MPP – 37,4% (рис. 4.2). При этом суммарная производительность совокупности из 500 наиболее производительных суперкомпьютеров должна была превысить 1,65 эксафлопс. В действительности в 2020 г. доля суперкомпьютеров архитектуры Cluster в общей производительности из списка Top 500 превысила 70%, а несколько меньшая, чем прогнозная, суммарная производительность (1,56 эксафлопс) была достигнута в 2019 г.

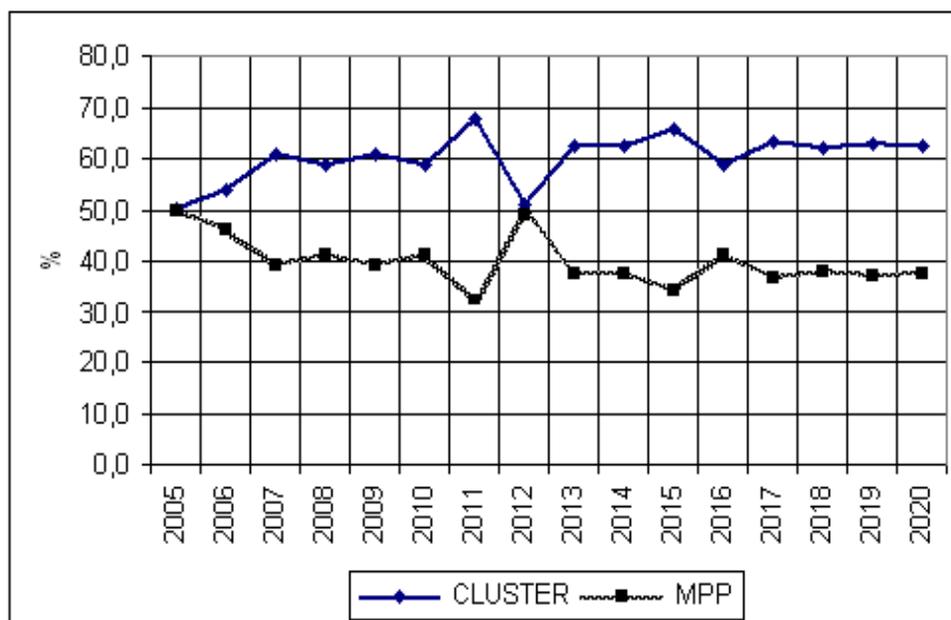


Рис. 4.2. Доля суперкомпьютеров с архитектурой MPP и Cluster в суммарной производительности всех суперкомпьютеров

4.2.4. Динамика показателей энергоэффективности суперкомпьютеров

Происходящее повышение производительности суперкомпьютеров, как правило, сопровождается ростом их электрической мощности. Так, суммарная электрическая мощность, потребляемая всеми суперкомпьютерами из списка Top 500 устойчиво возрастала со 161 МВт в 2007 г. до 696 МВт в 2016 г. (по данным ноябрьских выпусков Green 500) [8].

Несмотря на существенный прогресс в повышении энергоэффективности, энергоемкость высокопроизводительных суперкомпьютеров остается высокой. В частности, согласно рейтингу Top 500, проведенному в ноябре 2016 г., наиболее производительный в мире китайский суперкомпьютер Sunway TaihuLight 93,0 Пфлопс, потреблял около 15,37 МВт эл. мощности. Второй по производительности суперкомпьютер Tianhe-2 (MilkyWay-2) с быстродействием 33,9 Пфлопс имел электрическую мощность 17,81 МВт, а необходимая электрическая мощность для третьего по производительности суперкомпьютера Titan, имеющего быстродействие 17,59 Пфлопс, также составляла немалую величину – 8,2 МВт. [9].

Следует также отметить, что из-за высокой плотности размещения вычислительных мощностей в ограниченном пространстве, проблема отвода тепла в суперкомпьютерах стоит более остро, чем в центрах обработки данных (ЦОД). В настоящее время, по мнению специалистов, наилучшие характеристики, как и для ЦОД,

имеет технология водяного охлаждения, которая позволяет снизить показатель PUE^{viii} с 1,5–2 до 1,05–1,06 [10, 11].

В связи с высокой энергоемкостью суперкомпьютеров, начиная с 2007 г. регулярно публикуются списки Green 500, в которых проводится ранжировка наиболее производительных суперкомпьютеров по величине соотношения между их производительностью R_{\max} и используемой электрической мощностью. В соответствии с ноябрьским за 2016 г. списком Green 500, максимальная удельная производительность суперкомпьютеров составляла 9462,1 Мегафлопс/Вт. В то же время, как показывает анализ приведенных в этом списке данных, у 46% суперкомпьютеров из 500 самых мощных, показатель энергоэффективности составлял менее 10% от максимального уровня. Следует отметить, что по уровню энергоэффективности отечественные суперкомпьютеры долгие годы почти на порядок уступали лучшим зарубежным образцам, хотя в последнее время ситуация улучшается (в 2016 г. у нас был достигнут уровень 2056,6 Мегафлопс/Вт, что соответствовало 112 месту в списке Green 500) [8].

В связи с высокой энергоемкостью современных суперкомпьютеров, представляет интерес экономико-статистический анализ факторов, влияющих на их удельную производительность (т.е. производительность в расчете на 1 Вт электрической мощности). В настоящей главе такой анализ проведен в статике (применительно к суперкомпьютерам петафлопсной производительности $R_{\max} > 1$ Пфлопс) и в динамике (применительно к наиболее производительным в 2005–2014 гг. суперкомпьютерам).

Так, в табл. 4.4 представлены коэффициенты корреляции удельной производительности суперкомпьютеров и их основных технических характеристик ($R_{\max} > 1$ Пфлопс) с 2010 по 2016 г. Полученные данные показывают, что влияние на удельную производительность исследовавшихся факторов является существенно менее значимым, чем на производительность, и неоднонаправленным (табл. 4.4, ср. с табл. 4.2).

Вместе с тем, из физических соображений ясно, что в целом определяющее влияние на производительность и электрическую мощность суперкомпьютеров должны оказывать количество ядер и тактовая частота процессоров, а также топологический размер схем. Так, на основе выборки, составленной из наиболее производительных в 2016 г. суперкомпьютеров (с производительностью $R_{\max} > 1$ Пфлопс), была получена следующая регрессионная зависимость между удельной производительностью R_{\max}/W_e в Мфлопс/Вт и числом ядер на 1 Вт PR/W_e :

^{viii} PUE (англ. power usage effectiveness) – отношение полной подведенной мощности к реально затраченной на работу вычислительных систем).

$$\ln\left(\frac{R_{\max}}{W_e}\right) = 0,761 \ln\left(\frac{PR}{W_e}\right) + 4,078, \quad R^2 = 0,608 \quad (4.5)$$

(0,239) (0,057)

(включение большего числа факторов в эту и приводимые ниже регрессионные зависимости приводит к незначимым с позиций t -статистик дополнительным коэффициентам).

Таблица 4.4

Динамика коэффициентов корреляции удельной производительности суперкомпьютеров (Мегафлопс/Вт) и их основных технических характеристик ($R_{\max} > 1$ Пфлопс)

| | Удельная производительность Мфлопс/Вт | Число ядер на 1 Вт | Доля ядер-ускорителей (сопроцессоров) в общем количестве ядер | Внутренняя частота процессоров |
|------|--|--------------------|---|--------------------------------|
| 2010 | 1,000 | 0,660 | 0,279 | 0,589 |
| 2012 | 1,000 | 0,906 | -0,202 | -0,744 |
| 2014 | 1,000 | 0,648 | 0,309 | -0,402 |
| 2016 | 1,000 | 0,415 | 0,439 | -0,192 |

Для представительной выборки показателей суперкомпьютеров петафлопсного класса ($R_{\max} > 1$ Пфлопс), содержащихся в ноябрьских выпусках Top 500 за 2013–2015 гг., получены следующие регрессионные зависимости между удельной энергоемкостью 1 ядра процессора W_e/PR в Вт/ядро и тактовой частотой FR в МГц:

$$2013 \text{ г.: } \ln\left(\frac{W_e}{PR}\right) = 2,404 \ln(FR) - 22,852, \quad R^2 = 0,613; \quad (4.6a)$$

(0,354) (2,744)

$$2014 \text{ г.: } \ln\left(\frac{W_e}{PR}\right) = 2,751 \ln(FR) - 25,430, \quad R^2 = 0,448; \quad (4.6b)$$

(0,441) (3,425)

$$2015 \text{ г.: } \ln\left(\frac{W_e}{PR}\right) = 2,380 \ln(FR) - 15,798, \quad R^2 = 0,637. \quad (4.6c)$$

(0,268) (2,079)

Как следует из (4.6a), (4.6b), значения коэффициентов полученных зависимостей при логарифмах составляют 2,380 и 2,751, что согласуется с данными о том, что потребляемая процессорами электрическая мощность зависит от тактовой частоты в степени 2–2,5 [12]. По этой причине, в частности, в последние почти 20 лет одним из важных направлений повышения энергоэффективности суперкомпьютеров стало увеличение не тактовой частоты, а числа процессоров и ядер.

В то же время проведенный в работе [13] статистический анализ *динамических рядов* показывает, что для выборки из наиболее производительных суперкомпьютеров с процессорами на основе схем разного топологического размера,

показатель $R_{\max t}$ в млрд Мфлопс связан следующей функциональной зависимостью с числом ядер процессоров PR_t и топологическим размером чипов hp_t в нм:

$$\ln R_{\max t} = 10,940 + 0,852 \ln \left(\frac{PR_t}{hp_t^2} \right), \quad (4.7a)$$

а необходимая для работы суперкомпьютера электрическая мощность W_{et} (в кВт) оказывается связанной с отношением PR_t/hp_t зависимостью:

$$\ln W_{et} = 4,476 + 0,443 \ln \left(\frac{PR_t}{hp_t} \right). \quad (4.7b)$$

Из последних зависимостей следует, в частности, что рост удельной производительности R_{\max}/W_e наиболее производительных суперкомпьютеров был во многом связан с увеличением числа ядер процессоров и уменьшением топологического размера микросхем.

4.2.5. Моделирование «гонки» производительности суперкомпьютеров между отдельными странами

Особенностью предыдущего десятилетия стала утрата на продолжительное время единоличного лидерства США в производительности суперкомпьютеров. Так, в 2010 г. наиболее производительным суперкомпьютером из списка Top 500 стал китайский суперкомпьютер NUDT производительностью 2,566 Пфлопс. В 2011 г. на первое место вышел японский суперкомпьютер K computer производительностью 10,51 Пфлопс. После короткого периода лидерства США в 2012 г. (суперкомпьютер Titan производительностью 17,59 Пфлопс) на первое место опять вышли китайские суперкомпьютеры Tianhe-2 (MilkyWay-2, с 2013 по 2015 г. включительно), а с 2016 г. суперкомпьютер Sunway, производительность которого более, чем в 5 раз выше, чем у суперкомпьютера Titan. В результате в конце 2016 г. общая производительность суперкомпьютеров США из списка Top 500 по показателю R_{\max} была на 50 Пфлопс меньше, чем суммарная производительность суперкомпьютеров Китая и Японии (соответственно 228 Пфлопс против 278 Пфлопс). Лишь после 2018 г. Китай потерял лидерство в списке суперкомпьютеров Top 500, когда рекордсменами стали США (2018–2019 гг.) и Япония (2020–2021 гг.) с суперкомпьютерами Summit и Fugaku.

Следует отметить, что создание производительных суперкомпьютеров в Китае осуществляется в рамках пятилетнего плана, с целью преодоления отставания в уровне экономического развития от уровня развитых стран, а также с целью обеспечения собственных разработок в области высоких технологий. Последнее вызывает озабоченность в США, так как в стремлении к обладанию Китаем собственными

высокими технологиями там усматривается угроза повышения военного потенциала этой наиболее населенной страны мира [14]. Вместе с тем, нельзя исключать того, что в ближайшем будущем лидерство в области наиболее производительного суперкомпьютера опять на некоторое время перейдет к американцам, которые реализуют ряд проектов, связанных с созданием эксафлопсного суперкомпьютера (см. далее).

Таким образом, предыдущее десятилетие отмечено началом «гонки» производительности суперкомпьютеров между ведущими странами мира. В связи с этим, представляет практический интерес межстрановой анализ перспектив развития суперкомпьютерного потенциала.

С этой целью в настоящей работе построена модель «гонки» производительности суперкомпьютеров между США с одной стороны и Китаем и Японией с другой в форме векторной авторегрессии (VAR):

$$\begin{pmatrix} \ln(R_{\max t}^{USA}) \\ \ln(R_{\max t}^{China+Japan}) \end{pmatrix} = \begin{pmatrix} \lambda_1 & \mu_1 \\ \lambda_2 & \mu_2 \end{pmatrix} \begin{pmatrix} \ln(R_{\max t-1}^{USA}) \\ \ln(R_{\max t-1}^{China+Japan}) \end{pmatrix} + \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}, \quad (4.8)$$

где $R_{\max t}^{USA}$, $R_{\max t}^{China+Japan}$ – суммарная производительность совокупности суперкомпьютеров США, а также Китая и Японии, $\lambda_i, \mu_i, g_i, i = 1, 2$ – постоянные коэффициенты. Параметры построенной модели представлены в табл. 4.5.

Таблица 4.5

Оценки параметров λ_i, μ_i, g_i модели (4.8), $i = 1, 2$

| Совокупность суперкомпьютеров с архитектурой | λ_i | μ_i | g_i | Коэффициент детерминации R^2 |
|--|------------------|------------------|-------------------|--------------------------------|
| США | 0,664 (0,121) | 0,194 (0,078) | 0,913 (0,264) | 0,982 |
| Китай+ Япония | 1,368 (0,131) | 0,000 | -1,917 (0,550) | 0,893 |

Расчеты на основе модели (4.8) показывают, что при сохранении существующих тенденций, в 2020 г. суммарная производительность мощных суперкомпьютеров Китая и Японии списке Top 500 должна была превысить соответствующий показатель в США на 32% (в начале 2017 г. величина этого превышения составляла почти 22%) (рис. 4.3). Однако ввиду того, что Китай, опасаясь санкции США, перестал с 2021 г. представлять свои наиболее мощные суперкомпьютеры в список Top 500 (в частности, Sunway Pro с возможно эксафлопсной производительностью), официальное лидерство сохраняется за американскими системами [15].

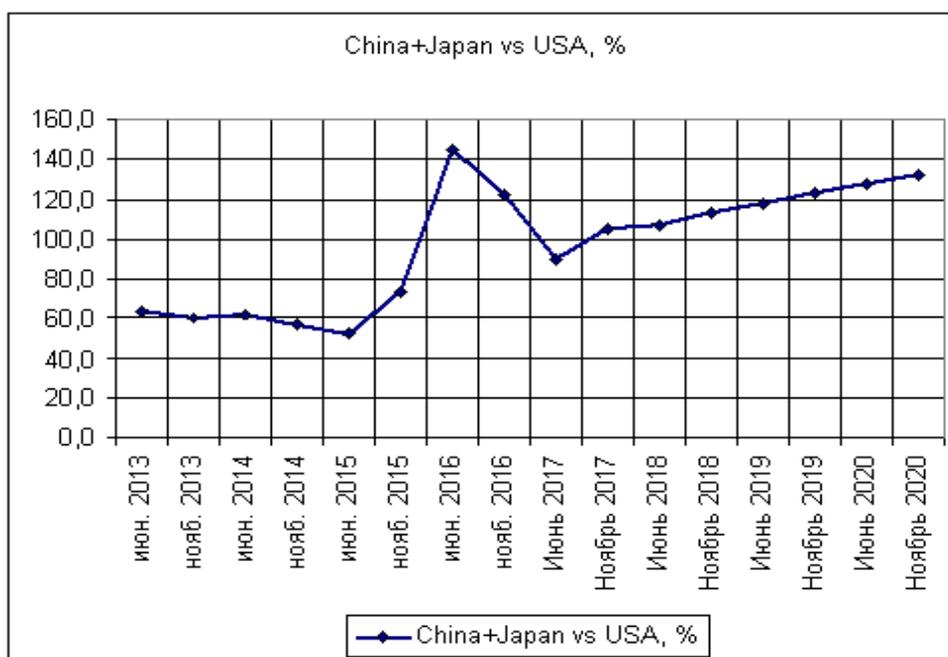


Рис. 4.3. Динамика соотношения между суммарной производительностью всех суперкомпьютеров Китая и Японии с одной стороны, и США с другой, %

4.4. Проблема создания суперкомпьютера экзафлопсной производительности

Бурный рост производительности суперкомпьютеров стимулировал возрастание спроса на высокопроизводительные вычисления НРС со стороны исследовательского сектора, государственных структур и бизнеса. В связи с этим, перед специалистами ряда высокоразвитых стран, в частности, в США, Японии и Китае, поставлена проблема создания суперкомпьютеров **экзафлопсной** производительности (т.е. производительностью более 10^{18} флоп в секунду). При этом необходимо иметь в виду, что в настоящее время не существуют принципиальных технических ограничений для создания суперкомпьютеров такого класса на основе существующего оборудования [16]. Однако, несмотря на достигнутые результаты в области энергоэффективности вычислительных устройств, даже при наилучшем в начале 2017 г. уровне энергоэффективности для устройств петафлопсного класса (9,462 Гигафлопс/Вт для суперкомпьютера DGX SaturnVc производительностью $R_{\max} = 3,307$ Пфлопс), необходимая электрическая мощность экзафлопсного суперкомпьютера составила бы 105,7 МВт. Этой мощности достаточно для обеспечения электроэнергией крупного городского квартала с населением несколько десятков тысяч человек. При полной загрузке его годовые расходы на электроэнергию составили бы порядка 70–80 млн долл (в условиях США).

Поэтому, например, в США, перед разработчиками была поставлена задача создания экзафлопсного суперкомпьютера электрической мощности порядка 20 МВт. Так, с целью ограничения роста масштабов энергопотребления, а также повышения его эффективности, агентством DARPA Министерства обороны США была в 2010 г. разработана программа «Вездесущие высокоэффективные вычисления» (Ubiquitous High Performance Computing – UHPC). В рамках этой программы ставилась задача обеспечить размещение в одном кабинете компьютерные мощности в 1 экзафлопс с энергоэффективностью 50 Гигафлопс/Вт (эта величина на порядок превышала текущий уровень). Предполагалось, что первый образец системы UHPC будет создан в 2018 г. [17].

Для первоначальной оценки реальности решения такой проблемы можно использовать зависимости (4.7a), (4.7b). Проведенные нами в 2013 г. расчеты с использованием этих зависимостей показывают, что при сохранении существующей тенденции увеличения производительности за счет использования большего числа ядер процессоров и уменьшения топологического размера схем до 10, 7 и 5 нм, мощность суперкомпьютера экзафлопной производительности может составить не 100 МВт, а соответственно 40, 33 и 29 МВт [13].

Вместе с тем, построенные зависимости не дают ответа на вопрос, когда может быть создан экзафлопсный суперкомпьютер с энергоэффективностью 50 Гигафлопс/Вт. Приближенную первоначальную оценку можно получить на основе исследования и экстраполяции сложившихся тенденций. С этой целью нами на основе временных рядов за 2005–2016 гг. построена авторегрессионная зависимость, связывающая логарифм максимальной в году t производительности (в гигафлопс) суперкомпьютеров из списка Top500 $\log(R_{\max 1t})$ с логарифмом соотношения между числом ядер процессоров PR_t и квадратом топологического размера чипов hp_t в нм $\log(PR_t/hp_t^2)$ (см. (4.7a):

$$\log(R_{\max 1t}) = \underset{(0,204)}{0,380} \log(R_{\max 1 t-1}) + \underset{(0,150)}{0,409} \log\left(\frac{PR_t}{hp_t^2}\right) + \underset{(0,950)}{3,148}, \quad R^2 = 0,966. \quad (4.9)$$

На основе (4.9) в 2016 г. был построен экстраполяционный прогноз максимальной производительности суперкомпьютеров $R_{\max 1t}$ на период до 2025 г. (рис. 4.4), при условии, что топологический размер схем будет уменьшаться каждые 2 года с масштабным фактором 0,7 (т.е. 10 нм в 2017 г., 7 нм в 2019 г. и т.д.). Тогда при сохранении действующих тенденций производительность в 1 экзафлопс могла бы быть достигнута в 2021 г.

Примерно к такому же выводу можно прийти при экстраполяции увеличения величины показателя энергоэффективности суперкомпьютеров на перспективу. Так,

при сохранении текущих тенденций энергоэффективность в 50 Гигафлопс/Вт могла бы быть обеспечена к 2021 г. Действительно, если бы она увеличивалась, как и в 2008–2016 гг. для наиболее энергоэффективных суперкомпьютеров петафлопсного класса ($R_{\max} > 1$ Пфлопс) со среднегодовым темпом прироста в 46,5%, то для достижения необходимой энергоэффективности в 50 Гигафлопс/Вт потребовалось бы несколько более 4 лет ($\ln(50/9,462)/\ln(1,465) = 4,4$ года).

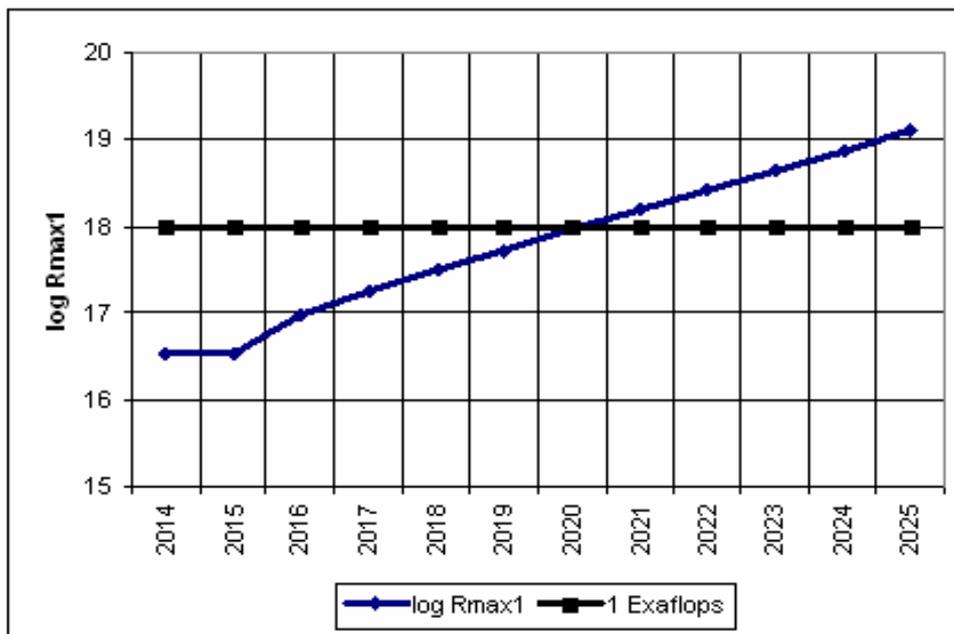


Рис. 4.4. Прогноз максимальной производительности суперкомпьютеров $R_{\max 1t}$ на основе (4.9)

Следует, однако, отметить, что в рамках статистического анализа не удается в полной мере отразить такие способы снижения энергоемкости будущего суперкомпьютера, которые связаны, в частности, с начавшимся недавно широким использованием сопроцессоров-ускорителей (малая длина временного ряда этого показателя не позволяет его использовать в статистической модели), а также с повышением производительности межсоединений. Так, в настоящее время одним из эффективных направлений увеличения производительности и снижения энергоемкости (повышения энергоэффективности) суперкомпьютеров является использование сопроцессоров (ускорителей) на основе высокопроизводительных графических процессоров – GPU (главным образом, GPU NVIDIA), а также сопроцессоров Intel Xeon Phi и ряда других производителей.

В целом полученные нами в [13] оценки свидетельствуют о реальности создания суперкомпьютера производительностью 1 эксафлопс и электрической мощностью 20–40 МВт при переходе на производство схем с малыми топологическими размерами. Возможные сроки такого перехода, сопряженного с высокими затратами,

зависят, в частности, от прогресса в создании: экономически эффективных установок экстремальной ультрафиолетовой фотолитографии (EUV); быстродействующих памяти и межсоединений; нового поколения массивно-параллельных систем; программного обеспечения, в частности, программных средств обнаружения и устранения ошибок, которые могут возникнуть ввиду повышения риска более частых неполадок из-за резкого увеличения (до миллиардов) числа компонентов в вычислительных системах (ensuring correctness in the face of faults; resilience at scale) и др. [18]. Одно из перспективных направлений ускорения перехода к энергоэффективным суперкомпьютерам состоит в создании мультиархитектурных вычислительных систем, в которых «...объединяются наборы процессоров для прикладных программ с различными формами параллелизма и набором функционально специализированных процессоров» [12].

В связи с дороговизной суперкомпьютеров существующих моделей и недостаточная эффективность этих устройств, построенных на базе компонентов общего назначения, для решения задач в области ИИ (прежде всего, машинного обучения). крупные производители микросхем приступили к разработке новых подходов к высокопроизводительным вычислениям на основе нейроморфных компьютеров и квантовых компьютеров. Одновременно исследуются перспективы использования устройств на основе графена, нанотрубок и др. продукции нанoeлектроники [19, 20].

Важно подчеркнуть, что в конечном счете перспективы развития суперкомпьютеров связаны с успешным решением проблемы как финансирования, так и снижения стоимости. Роль показателя стоимости особенно высока ввиду как дороговизны высокопроизводительных суперкомпьютеров, так и относительно малой продолжительности их жизненного цикла по сравнению с другими высокотехнологичными объектами сопоставимой стоимости. Так, из 7 эксплуатировавшихся в 2010 г. высокопроизводительных суперкомпьютеров ($R_{\max} > 1$ Пфлопс), к 2016 г. в списках Top 500 и Green 500 осталось всего 3. Наглядным примером может служить судьба суперкомпьютера Roadrunner в Лос-Аламосской национальной лаборатории, который первым превзошел петафлопсный рубеж производительности в 2008 г. Эксплуатация этого суперкомпьютера стоимостью 121 млн долл. была прекращена в конце марта 2013 г., т.е. менее, чем через 5 лет после ввода в строй. [21].

Стоимость введенных в эксплуатацию несколькими годами позже, а также разрабатываемых высокопроизводительных машин на порядок выше, чем у Roadrunner. Так, по некоторым данным стоимость первого в списке Top 500 в ноябре 2011 г. японского суперкомпьютера K Computer превысила миллиард долларов [16]. По некоторым данным, стоимость одной экзафлопсной системы, разрабатываемой одновременно в США и Китае, составит 300–500 млн долл. Кроме того, до ввода

первого образца потребуется ежегодно не менее 1 млрд долл. для проведения необходимых исследований и разработок. Желаемый объем финансирования суперкомпьютера нового поколения экзафлопсной производительности, который намечалось создать в Японии, также составлял порядка 1 млрд долл. (11 млрд иен) [22–24].

Вместе с тем, в более позднем варианте Проекта создания экзафлопсного суперкомпьютера в США (U.S. Exascale Computing Project) было предусмотрено создание первого такого американского суперкомпьютера в 2021 г. В Китае первый экземпляр экзафлопсного суперкомпьютера предполагалось создать в 2020 г. [24].

Следует отметить, что реальные сроки создания экзафлопсных систем оказались близкими к оценкам, полученным нами в 2016 г. [5]. Так, в мае текущего 2022 г. было объявлено о начале новой, экзафлопсной эры в списке Top 500. Возглавил этот список суперкомпьютер Frontier производительностью в 1,102 экзафлопс и электрической мощностью 21,1 МВт [25]!

4.5. Заключение

Разработанные в главе модели позволяют провести анализ динамики распространения высокопроизводительных суперкомпьютеров, их производительности, изменения структуры, а также факторов, влияющих на энергоэффективность. Они могут быть также использованы при первоначальной оценке перспектив достижения устанавливаемых на будущее уровней производительности и эффективности суперкомпьютеров.

Полученные нами оценки позволили подтвердить тезис о реальности создания суперкомпьютера производительностью 1 экзафлопс и электрической мощностью 20–40 МВт при переходе на производство схем с меньшими топологическими размерами.

Список литературы к главе 4

1. URL: <https://www.hpcwire.com/2017/04/06/supercomputer-sales-drove-2016-hpc-market-record-11-2-billion/>
2. URL: <http://www.top500.org>
3. Turek D. The Age of the Supercomputer. URL: <http://www.businessweek.com/stories/2007-05-07/the-age-of-the-supercomputerbusinessweek-business-news-stock-market-and-financial-advice>
4. Левин В.К. Высокопроизводительные вычислительные системы – суперкомпьютеры в России // Базовые лекции по электронике. Т. II. Твердотельная электроника / под общ. ред. В.М. Пролейко. М.: Техносфера, 2009. С. 576–594.

5. Варшавский Л.Е. Экономико-математическое моделирование тенденций изменения ключевых показателей высокопроизводительных суперкомпьютеров // Труды Института системного анализа Российской академии наук. 2017. Т. 67. № 3. С. 64–74.
6. Варшавский Л.Е. Анализ и моделирование динамики ключевых показателей высокопроизводительных суперкомпьютеров // Концепции. 2015. № 2 (34). С. 22–35.
7. Сахал Д. Технический прогресс: концепции, модели, оценки. М.: Финансы и статистика 1985.
8. URL: <http://www.green500.org/lists.php>
9. URL: <http://www.cnews.ru/news/top/index.shtml?2012/06/27/494386>
10. Левшин И. Из жизни алгоритмов. HPC Superstar // Суперкомпьютеры. 2013. № 13. С. 10–12.
11. URL: <http://efytimes.com/e1/99599/Eurotech-Supercomputer-At-CINECA-Breaks-World-Record-For-Energy-Efficiency>
12. Митропольский Ю.И. Элементная база и архитектура будущих суперкомпьютеров // Микроэлектроника. 2015. Т. 44. № 3. С. 163–179.
13. Варшавский Л.Е. Проблемы повышения энергоэффективности аппаратных средств в области информационных технологий // Труды Института системного анализа Российской академии наук. 2013. Т. 63. Вып. 3. С. 3–19.
14. Feldman M. The DOE and NSA Construct Doomsday Scenario for American HPC. URL: <https://www.top500.org/news/the-doe-and-nsa-construct-doomsday-scenario-for-american-hpc/>
15. URL: <https://www.hpcwire.com/2022/05/30/hyperion-hpc-market-is-stabilizing-and-headed-to-50b-by-2026> (дата обращения 01.06.2022).
16. Левшин И. Из жизни звезд // Суперкомпьютеры. 2013. № 13. С. 6–9.
17. Feldman M. DARPA Sets Ubiquitous HPC Program in Motion. URL: http://www.hpcwire.com/hpcwire/2010-08-10/darpa_sets_ubiquitous_hpc_program_in_motion.html
18. Reed D., Dongarra J. Exascale Computing and Big Data // Communications of the ACM. 2015. Vol. 58. No. 7. Pp. 56–68.
19. URL: <https://www.hpcwire.com/2018/01/25/hpc-ai-two-communities-future/>
20. Варшавский Л.Е. Проблемы спроса и предложения (на примере отдельных областей нанoeлектроники) // Экономические проблемы развития революционных технологий. Нанотехнологии: монография / рук. авт. коллектива В.Л. Макаров, А.Е. Варшавский. Гл. 7. М.: Наука, 2012. С. 132–157.
21. URL: <http://www.bbc.co.uk/news/technology-21993132>
22. URL: <http://efytimes.com/e1/fullnewsp.asp?edid=105870>
23. URL: <http://fcw.com/articles/2013/07/17/exoflop-supercomputing.aspx>
24. URL: <https://www.hpcwire.com/2017/04/26/messina-update-u-s-path-exascale-15-slides/>
25. URL: <https://www.hpcwire.com/2022/05/30/the-final-frontier-us-has-its-first-exascale:super-computer/> (дата обращения 01.06.2022).

ГЛАВА 5.

ПРОБЛЕМЫ РАЗВИТИЯ ИНДУСТРИИ КВАНТОВЫХ ВЫЧИСЛЕНИЙ

5.1. Введение

Последнее десятилетие характеризуется интенсификацией работ по подготовке к практическому использованию квантовых информационных систем. Получены серьезные результаты в области технологии квантовых вычислений и связи. Все большее число компаний начинает тестировать квантовые компьютеры и приборы с целью оценки возможностей их использования в будущем. Началась гонка квантовых технологий, в которую оказываются вовлеченными крупные компании и страны. Вместе с тем, в индустрии квантовых информационных технологий сохраняются проблемы принципиального характера, которые должны быть решены для реализации широкомасштабного использования этих технологий.

В данной главе проводится инженерно-экономический анализ проблем, связанных с индустрией квантовых вычислений. Материал главы основан на статье автора [1].

5.2. Особенности квантовых компьютеров

В отличие от обычного классического компьютера, оперирующего с битами, квантовый компьютер оперирует с состояниями квантовых систем. Простейшей квантовой системой, которая может выполнять функцию, аналогичную битам в классических компьютерах, является система с двумя возможными состояниями – кубит (qubit) – квантовый бит информации. Любое нормированное состояние однокубитной квантовой системы ψ , является суперпозицией двух базисных состояний $|0\rangle$ и $|1\rangle$:

$$|\psi\rangle = a|0\rangle + b|1\rangle, \text{ где } |a|^2 + |b|^2 = 1.$$

Измерение состояния кубита приводит к одному из базисных состояний $|0\rangle$ и $|1\rangle$ с вероятностями соответственно $|a|^2$ и $|b|^2$. Суперпозиция состояний квантовой системы, состоящей из n кубитов, также является состоянием квантовой системы. В связи с возможностью суперпозиции состояний, квантовый компьютер с n кубитами позволяет проводить вычисления в гильбертовом пространстве размерностью 2^n , что резко повышает его вычислительные мощности (при $n = 1000$ размерность пространства составляет $2^{1000} \approx 10^{300}$) [1, 2].

Другим важным свойством является такое явление, как запутанность (entanglement) состояний квантовых систем, характеризующее коррелированность состояний квантовых систем (кубитов). Использование запутанности позволяет повысить эффективность квантовых вычислений, обеспечить надежную связь (включая телепортацию) и создать высокоточные приборы (датчики электромагнитного излучения высокого разрешения и др.) [3, 4].

Логические операции над квантовыми состояниями в квантовых компьютерах выполняются путем унитарных преобразований, не нарушающих квантовые суперпозиции в процессе вычислений. Унитарные преобразования (квантовые гейты), по сути, представляют собой логические блоки (вентили), составляющие процесс вычислений. Для реализации любой унитарной операции на n кубитах достаточно иметь универсальный конечный набор однокубитовых операций (например, операции НЕ, преобразования Адамара и др.) и двухкубитовой операции CNOT (условного НЕ) [2, 5, 6].

Схематично работу квантового компьютера можно представить в виде последовательности трех операций: приведение системы в начальное состояние; проведение вычислений путем унитарных преобразований начальных условий; измерение конечного состояния) [7].

Особенность квантового компьютера состоит в том, что правильное решение в нем может быть определено лишь с некоторой вероятностью. Специалистами отмечается, что разработанные к настоящему времени алгоритмы для квантовых вычислений приводят, как правило, к правильным результатам с достаточно большой вероятностью за небольшое число итераций [2].

Повышенный интерес к разработке квантового компьютера, наблюдаемый в последние годы, обусловлен тем, что в 1990-х гг. в Bell Laboratories были разработаны эффективные алгоритмы квантовых вычислений, позволяющие в принципе решить ряд важных задач, для которых не существует эффективных алгоритмов решения на традиционных компьютерах (например, имеющую важное значение для криптографии задачу о факторизации чисел, проблему коммивояжера и связанные с ней задачи исследования операций и др.). Важным стимулом к повышению интереса к квантовым компьютерам и квантовым вычислениям явилась разработка Гровером (Grover) в 1995 г алгоритма поиска в неструктурированной базе. Квантовый алгоритм Гровера позволяет существенно снизить число итераций, необходимых для поиска N случайно заданных объектов (с $N/2$ до $N^{1/2}$).

Достигнутые в последние годы успехи в области создания квантовых алгоритмов позволили начать использование квантовых компьютеров в облачных вычислениях, например, для поисково-информационных систем. Реализация квантовых

алгоритмов на квантовых компьютерах позволила бы, по некоторым оценкам, существенно снизить парк серверов, а, следовательно, и капитальные и текущие затраты (включая затраты на энергию) в центрах обработки данных (ЦОД) [8].

Другой перспективной областью может быть применение квантовых компьютеров для моделирования химических процессов. Следует учитывать, что во многих практически важных случаях исследование квантовых физико-химических систем пока еще не удастся провести даже на современных суперкомпьютерах [4].

5.3. Некоторые подходы к созданию квантовых компьютеров

К настоящему времени можно выделить 2 основных наиболее продвинутых направления исследований по созданию квантовых компьютеров: 1) на основе *сверхпроводниковых кубитов*; 2) с кубитами на основе *ионных ловушек*.

За последнее десятилетие наметился определенный прогресс в создании кубитов в металлах, находящихся в сверхпроводящем состоянии (в таких кубитах в качестве металла обычно используется ниобий или алюминий). Работа сверхпроводящих кубитов основана на использовании джозефсоновских переходов. Перспективы этой технологии связаны с тем, что переключение логических схем в сверхпроводящем состоянии (путем изменения интенсивности и направления магнитных полей) осуществляется с существенно меньшим выделением энергии, чем в кремниевых чипах. Кроме того, несмотря на необходимость использования криогенной техники, технологии охлаждения материалов до температуры жидкого гелия уже хорошо освоены [8].

В настоящее время существуют две основные разновидности архитектуры квантовых сверхпроводящих компьютеров. Первый тип компьютеров связан с использованием и управлением квантовыми гейтами что представляет собой сложную задачу, учитывая то, что для решения практических задач требуется большое число кубитов (сотни и тысячи). Компьютеры этого типа считаются универсальными, т.е. направленными на решение широкого круга задач. Именно такого типа сверхпроводниковые квантовые компьютерные чипы разрабатываются в таких компаниях-гигантах, как Intel, IBM, Google.

В последнее время сотрудники специализированных центров этих компаний заняты созданием и тестированием прототипов сверхпроводящих квантовых чипов, состоящих примерно из 50–70 кубитов [9]. На сверхмощных суперкомпьютерах проводятся имитационные эксперименты с целью проверки возможностей таких чипов, в частности проверки достижения квантового превосходства (quantum supremacy). В январе 2019 г. IBM объявила о создании первой в мире квантовой системы,

которая может быть использована в коммерческих целях – Q System One, включающей 20 кубитов [10].

Значительный прогресс достигнут в создании сверхпроводящих квантовых компьютеров второго типа – *адиабатических квантовых компьютеров* (AQC), разработку которых осуществляет канадский стартап D-Wave Systems. Особенностью адиабатических (в англоязычной литературе часто используется термин «annealing») квантовых компьютеров является то, что в них не используются квантовые гейты. Эти компьютеры можно отнести к классу аналоговых, которые предназначены для решения специализированных задач.

Адиабатические квантовые вычисления основаны на использовании адиабатической теоремы квантовой механики, вытекающей из уравнения Шредингера. Суть этой теоремы состоит в то, что при медленном изменении (или воздействии), квантовая система, находящаяся в начальный момент времени в минимальном энергетическом исходном состоянии, в конечном итоге перейдет в минимальное энергетическое состояние измененной системы.

Математически это можно вкратце объяснить следующим образом. Нестационарное уравнение Шредингера имеет вид:

$$i\hbar\frac{\partial}{\partial t}\Psi(t) = H(t)\Psi(t), \quad (5.1)$$

где $H(t)$ – гамильтониан квантовой системы; $\Psi(t)$ – волновая функция, которая может быть представлена в виде ряда:

$$\Psi(t) = \sum_n c_n(t)\psi_n(t)e^{i\theta_n(t)}, \quad (5.2)$$

где $\psi_n(t)$, $n = 1, 2, \dots$ – ортонормированная система собственных функций оператора $H(t)$ (здесь и далее для удобства записи в скобках не представлена переменная положения); $\theta_n(t) = \int_0^t E_n(\tau)d\tau$ – динамическая фаза; $E_n(t)$ – уровень энергии, соответствующий собственной функции $\psi_n(t)$ (собственное значение оператора $H(t)$); i – мнимая единица. Можно показать, что в силу ортонормированности $\psi_n(t)$, $n=1,2,\dots$, динамика $c_n(t)$ связана со значениями всех остальных коэффициентов $c_m(t)$, $m \neq n$ и производной гамильтониана $\frac{d}{dt}H(t)$. Но если $\frac{d}{dt}H(t) \approx 0$ (*условие адиабатичности*) $c_n(0) = 1$, а все остальные $c_m(0) = 0$, $m \neq n$, то можно показать, что

$$\Psi(t) = c_n(t)\psi_n(t)e^{i\theta_n(t)} = \psi_n(t)e^{i\theta_n(t)}e^{i\gamma_n(t)}, \quad (5.3)$$

где $\gamma_n(t)$, так называемая, геометрическая фаза.

Иными словами, квантовая система, первоначально находившаяся на n -м энергетическом уровне, останется на том же порядковом энергетическом уровне измененной системы, но, естественно с другой величиной энергии (собственного значения $E_n(t)$ гамильтониана $H(t)$ в момент t) (более подробно см. например, статью: Adiabatic Theorem, Wikipedia).

Таким образом, решение оптимизационных задач осуществляется в следующей последовательности. Вначале формируется гамильтониан системы (с взаимодействующими кубитами), минимальное энергетическое состояние которой (ground state) может быть решением оптимизационной проблемы (гамильтониан конечной системы). На следующем шаге синтезируется система с простым гамильтонианом, для которого может быть достаточно просто определено минимальное энергетическое состояние (например, система с невзаимодействующими кубитами). В это начальное состояние переводится синтезированная квантовая система. Далее, простой гамильтониан медленно (т.е. адиабатически в квантово-механическом смысле) преобразуется в гамильтониан конечной системы. При этом, в соответствии с адиабатической теоремой, начальное минимальное энергетическое состояние простой системы переводится в минимальное энергетическое состояние для конечной системы, гамильтониан которой соответствует функционалу решаемой оптимизационной задачи. Полученное состояние и является решением задачи [8].

В 2011 г. состоялась первая продажа квантового компьютера D-Wave One, созданного D-Wave Systems, аэрокосмической компании Lockheed Martin за 10 млн долл. Компьютер из 128 кубитов, находящихся в сверхпроводящем состоянии, предназначался для решения специализированных задач, связанных с созданием сложных радиолокационных и аэрокосмических систем. В частности, предполагалось использовать D-Wave One для проведения исследований по автоматическому обнаружению ошибок в программных продуктах для управления сложными системами (в рамках проекта создания истребителя F-35). Уже после двухлетней эксплуатации специалисты компании Lockheed Martin стали проявлять больший оптимизм в отношении возможностей коммерческого использования приобретенного компьютера [11, 12].

В 2013 г. NASA, Google и консорциум университетов (Universities Space Research Association – USRA) организовали Лабораторию квантового искусственного интеллекта (Quantum Artificial Intelligence Lab), в которой стали проводить исследование возможностей адиабатического квантовых компьютеров D-Wave, начиная с D-Wave Two, созданного путем апгрейда первого коммерческого квантового компьютера D-Wave One со 128 до 512 кубитов. Система D-Wave Two использовалась в исследованиях по машинному обучению, при моделировании атмосферных и

др. явлений (в NASA), при создании системы поиска в Интернете и технологии распознавания речи (в Google) [13]. Принятию решения об учреждении Лаборатории предшествовало тестирование D-Wave Two независимыми экспертами и всеми сторонами-учредителями. Оно показало, что с помощью адиабатического квантового компьютера ряд проблем может быть выполнен на 3 порядка быстрее, чем с помощью обычного.

В 2017 г. D-Wave выпустила систему 2000Q со свободным программным обеспечением. Система состоит из 2048 сверхпроводящих кубитов, в ней используется 6016 соединений (couplers) и 128000 джозефсоновских переходов. Рабочая температура системы – 15 мК (15 миллиКельвин, что ниже, чем в квантовых системах Intel, IBM, Google)!

Недавно D-Wave запустила систему виртуализации квантовых облачных вычислений Leap, с предоставлением широкой аудитории пользователей возможности бесплатного использования своей системы 2000Q в «облаке» в течение 1 минуты в месяц. Для проведения облачных вычислений стартап предлагает свободное программное обеспечение (open-source software development kit – SDK) под названием Ocean. В то же время для пользователей, заинтересованных в более длительной работе с системой, установлен тариф в 2000 долл. за час в месяц [14].

В 2019 г. D-Wave объявила о работе по созданию системы из 5000 кубитов, которую намечалось завершить в середине 2020 г. В создававшемся новом чипе Regasus каждый кубит должен был иметь 15 соединений с другими кубитами.

Сообщалось также о контракте в 100 тыс. долл., подписанном Ford Motor Company с Лабораторией квантового искусственного интеллекта, дающей возможность автомобильной компании использовать для решения задачи коммивояжера квантовый компьютер D-Wave 2000Q, установленный в NASA. Отмечается также, что еще ранее, в 2017 г. другой автомобильный гигант, Volkswagen, использовал квантовый компьютер D-Wave для оптимизации маршрутов 10 тыс. такси в Пекине. Вместе с тем, по мнению специалистов, все еще открытым остается вопрос о преимуществах использования систем D-Wave, по сравнению с традиционными компьютерами, для решения данной задачи [15].

Об интересе, проявляемому к возможностям квантовых компьютеров в транспортной отрасли, свидетельствует их эксплуатация в таких компаниях, как Daimler, BMW, Ford, Toyota, Airbus. В ноябре 2018 г. Volkswagen провел серию расчетов на квантовом компьютере с целью более точного прогнозирования дорожного движения и спроса на транспортные услуги. По убеждению сотрудников компании, проведенные с использованием квантовых компьютеров расчеты позволяют

заблаговременно с высокой вероятностью (95%) определить, где должно будет находиться такси через час [16].

Airbus использует квантовые компьютеры D-Wave для оптимизации взлета и посадки самолетов на основе обработки большого объема данных с многочисленных датчиков (число их на современных самолетах компании достигает нескольких сотен тысяч), для более точных аэродинамических расчетов, а также для улучшения конструкций крыла.

Другим свидетельством возможностей D-Wave 2000Q является сообщение о мировом рекорде, установленном в текущем году китайскими учеными в проблеме факторизации. С использованием 89 кубитов этого квантового компьютера им удалось представить 20-битное число 1 005 973 в виде произведения простых чисел, т.е. в виде $1\ 005\ 973 = 997 \cdot 1009$ (ранее удавалось факторизовать в 3 раза меньшее число: 376 289). Полученный результат является еще одним шагом по пути дешифрования широко используемой в цифровой экономике криптографической системы RSA-2048 (основывается на факторизации больших чисел), хотя для факторизации 2048-битных чисел, используемых в этой системе потребуются, по оценкам, 10 тыс. кубитов адиабатического квантового компьютера [17].

Следует отметить, что прообразы простейших квантовых процессоров на основе сверхпроводящих кубитов, для которых получены запутанные состояния, а также более высокое, чем ранее, время поддержания когерентности, созданы и в ряде университетских лабораторий (см., например, [18–22]).

Определенные успехи достигнуты и в создании квантовых компьютерных систем с кубитами на основе *ионных ловушек*. Так, в университете штата Мэриленд (США) под руководством С. Монрое и в Гарвардском университете под руководством М. Лукина, приблизительно в одно время (в 2017 г.) созданы квантовые компьютеры, состоящие из более, чем 50 кубитов. Такого вида кубиты отличаются высоким качеством гейтов, характеризуемым низкой чувствительностью к помехам (ошибка не превышает нескольких единиц на миллион операций гейта в случае однокубита). Однако отмечаются трудности в создании монолитных квантовых систем с более, чем 50 кубитами. По этой причине предлагается исследовать модульную конструкцию квантовых компьютеров данного вида [23, 24].

5.4. Гонка за лидерство в области квантовых технологий

Успехи, достигнутые в области квантовых технологий, не ограничиваются только квантовыми вычислениями. Значительное внимание в развитых странах уделяется и проблеме создания надежных квантовых коммуникаций. Так, в Китае еще

в 2016 г., с целью исследования технологии квантового шифрования, механизма квантовой запутанности, а также тестирования квантовой телепортации, был запущен спутник квантовой связи «Мо-Цзы». В ходе одного из экспериментов, передавались данные со спутника «Мо-Цзы» на наземные наблюдательные станции, причем расстояние между спутником и наземными станциями составляло от 645 до 1200 км. Применение технологии квантового шифрования позволило существенно повысить скорость передачи данных по сравнению с условной оптоволоконной связью из космоса. По мнению ряда экспертов, уже к 2030 г. Китай сможет создать невзламываемую конфиденциальную сеть, устойчивую к любым атакам. Сообщается также о создании в Китае квантового радара, который сможет распознавать невидимые для обычных РЛС самолеты «Stealth», а также подводные лодки.

Наконец, в 2019 г. в Китае запущен беспилотник, который может быть использован в качестве промежуточного узла спутниковой квантовой связи и служить мобильной, менее затратной альтернативой оптоволоконной связи [25].

Ведутся работы в области создания квантового интернета. В начале 2019 г. в Британии была введена в строй первая в мире коммерческая квантовая сеть, в которой для защиты зашифрованных сообщений используются принципы квантовой механики.

Таким образом, квантовые информационные технологии превратились в технологии двойного назначения, уровень развития которых может определять военное и экономическое превосходство стран. В связи с этим, в ведущих экономических державах мира разработаны планы и программы развития квантовых технологий. Так, по оценкам ряда экспертов, Китай ежегодно расходует миллиарды долларов на работы в области квантовых технологий. В частности, в ближайшие 10 лет в этой стране намечено выделить 10 млрд долл. на развитие квантовой индустрии, из которых 3 млрд долл. – на развитие квантовых вычислений.

В результате за последние годы Китай стал лидером по числу поданных патентов в области квантовых технологий. В 2018 г. число патентов в этой области, зарегистрированных китайскими специалистами, составило 492, что вдвое превышало число соответствующих патентов в США. Большая часть китайских патентов относится к квантовой коммуникации. Вместе с тем, по числу патентов в области квантовых вычислений (193 в 2018 г.) США продолжает уверенно лидировать (у Китая их было 63) [26]. В то же время, по числу публикаций в области квантовых вычислений за период с 2013 г. лидером является Китай (2986 против 2494 публикаций в США) [27].

В самом конце 2018 г. Президентом США был подписан Закон о национальной квантовой инициативе (National Quantum Initiative Act). На развитие квантовых информационных технологий выделяется 1,2 млрд долл. в течение 10 лет. В этом

мероприятия будут участвовать как государственные организации (NASA), Национальный институт стандартов и технологии (National Institute of Standards and Technology), так и промышленность и университеты. В рамках принятой инициативы Национальный научный фонд США и Министерство энергетики США должны будут создать от двух до пяти исследовательских и образовательных центров в области квантовых информационных технологий [28].

В Великобритании с 2014 г. реализуется Национальная программа развития квантовых технологий (The National Quantum Technologies Programme). В ближайшие годы (с 2019 до 2024 г.) на нее предполагается выделить 315 млн фунтов стерлингов. В рамках программы ведется исследование и разработка квантовых датчиков, квантовых часов, квантовых систем навигации и связи, квантовых вычислений и др. направлений [29].

Программы в области развития квантовых технологий приняты и в Южной Корее, Израиле, на Тайване и в других странах, хотя на них выделяются значительно меньшие финансовые средства.

В России также осуществляются интенсивные исследования в области квантовых технологий. Так, в начале 2018 г. Внешэкономбанком и рядом исследовательских центров («ВЭБ-Инновации», Фонд перспективных исследований, МГУ имени М.В. Ломоносова, АНО «Цифровая экономика») было подписано соглашение, связанное с созданием в стране многокубитного оптического квантового симулятора на основе фотонных чипов и нейтральных атомов, состоящего из не менее пятидесяти кубитов. В ходе работ по проекту уже выполнен ряд экспериментов по созданию ловушек для массивов нейтральных холодных атомов. Предполагалось, что квантовый компьютер будет создан к 2021 г.

В марте 2019 г. в Российском квантовом центре (РКЦ) и НИТУ МИСиС был разработан план развития квантовых технологий, в соответствии с которым в 2024 г. страна должна была довести свою долю на глобальном рынке квантовых коммуникаций до 8%. Реализация этого плана потребует 44 млрд рублей, причем половина этих средств, по мнению разработчиков, должна быть направлена на гранты, ориентированные на проекты по разработке и применению технологий [30].

5.5. Принципиальные проблемы, требующие решения для практического использования квантовых вычислений

Создание квантовых компьютеров сдерживается необходимостью решения ряда принципиальных проблем. Одна из главнейших проблем связана с необходимостью преодоления эффектов декогерентизации, которые возникают при

взаимодействии квантового компьютера с окружающей средой, при неточности задания параметров управляющих сигналов или вследствие неконтролируемого взаимодействия между собой кубитов и т.д. (декогерентизация препятствует проявлению суперпозиции и запутывания состояний, то есть, использованию именно тех свойств квантового компьютера, которые обуславливают его преимущество по сравнению с традиционным). Для решения этой проблемы в настоящее время разрабатываются различные варианты метода **квантового исправления ошибок (quantum error correction)** [2]. Однако в этом направлении существуют практические трудности, так как для борьбы с ошибками и помехами, по расчетам специалистов, потребуется 1000 исправляющих (error correcting) кубитов на один вычислительный кубит. По этой причине считается, что в течение ближайших 5–10 лет будут создаваться квантовые компьютеры без использования полноценных методов исправления ошибок, т.е. так называемые зашумленные квантовые компьютеры промежуточного масштаба – NISQ (Noisy Intermediate-Scale Quantum) [23].

В настоящее время, в связи с малостью времени когерентности (так, в квантовых компьютерах IBM время когерентности составляет примерно 70 микросекунд), а также накоплением ошибок в цепи вычислений (средняя частота ошибок однокубитных гейтов в этих компьютерах составляет 0,041%), разработчики создают **цепи с малой глубиной (short-depth circuits)**, т.е. с малым числом операций, при которых сохраняется когерентность [31, 32].

Кроме того, измерение кубитов сопряжено с решением сложных технологических проблем, требующих преодоления трудностей измерения состояния отдельной атомной частицы: атома, иона, электрона, спина электрона или атомного ядра, фотона.

Ввиду отмеченных выше проблем, ряд специалистов пришел к выводу о невозможности создания полноценного квантового компьютера в обозримом будущем. Так, известный физик М. Дьяконов, работающий ныне во Франции, в своей интересной статье так объясняет свой вывод. Для того, чтобы конкурировать с современным ноутбуком (laptop), квантовый компьютер должен иметь от 1000 до 100 тыс. кубитов. Такой компьютер должен непрерывно оперировать с $2^{1000} = 10^{300}$ параметрами, что превышает число субатомных частиц во Вселенной. Практически нерешаемой является и проблема исправления ошибок. Так, при выполнении условий **пороговой теоремы (threshold theorem)**, заключающихся, в частности, в том, что частота ошибки в расчете на один кубит и один гейт не должна превышать некоторого предельного уровня, теоретически становится возможной бесконечное число квантовых вычислительных операций, но за счет добавления новых, *логических* кубитов. Однако добавление тысяч и сотен тысяч новых кубитов может только

усугубить проблему. М. Дьяконов отмечает также проблему измерения большого числа непрерывных значений параметров, указывая на трудности точного определения даже квадратного корня из 2 [33].

Определенный скептицизм относительно возможностей коммерческого использования NISQ компьютеров выражен в докладе Национальной академии наук США (National Academies of Science, Engineering and Medicine, сокр. NAS). Так, в нем отмечается, что в настоящее время не известно об алгоритмах или приложениях, которые могли бы быть эффективно использованы в этих компьютерах [34].

В то же время, оптимистичные разработчики убеждены, что после успешного решения отмеченных проблем, квантовые компьютеры найдут широкое применение через 5–10 или 12 лет [35].

Таким образом, в целом, работы по созданию *универсальных* квантовых компьютеров находятся в стадии фундаментальных исследований и еще далеки от перехода в стадию внедрения в производство. Несмотря на проявляемый в последние месяцы со стороны ряда исследователей оптимизм, разработка необходимых для решения широкого круга задач квантовых компьютеров потребует достаточно длительного времени, причем для получения положительного результата необходима консолидация усилий в нанoeлектронике, других областях физики, а также в математике. [2, 4, 36].

5.6. Экосистема индустрии квантовых вычислений

В настоящее время, наряду с компаниями-гигантами, активное участие в создании индустрии квантовых вычислений принимают и стартапы. Кроме упоминавшегося D-wave, наиболее крупными являются Rigetti Computing, PsiQ, Silicon Quantum Computing, Cambridge Quantum Computing Ltd (CQCL). С 2012 г. эти стартапы получили соответственно 205 млн, 119 млн, 65 млн, 60 млн и 50 млн долл. (по состоянию на конец 2018 г. [23]). В целом же, все стартапы, связанные с этой тематикой, получили за указанный период 700 млн долл.

Поддержку стартапам оказывают крупные промышленные и финансовые компании. Интересно отметить, что инвесторами стартапа D-Wave являются Jeff Bezos (основатель Amazon.com), инвестиционный банк Goldman Sachs и инвестиционная фирма In-Q-Tel, имеющая тесные связи со спецструктурами США. Подобный интерес к квантовым компьютерам со стороны бизнес-сообщества во многом обусловлен тем, что по мере приближения к пределам кремниевой электроники уже завершается эра экспоненциального роста производительности классических компьютеров [37, 38].

8 стартапов (QC, Zapata Computing и др.) входят в сеть IBM Q Network, организованную компанией IBM. Однако ввиду того, что, по мнению многих специалистов, превосходство квантовых компьютеров над традиционными для решения для решения практических задач будет достигнуто не в самой ближайшей перспективе, венчурный капитал в целом не торопится вкладывать значительные суммы в поддержку квантовых стартапов [39].

Тесно сотрудничают с промышленными и финансовыми компаниями-гигантами и компании ИТ-отрасли. Так, в целях развития экосистемы квантовых вычислений IBM объявила о создании консорциума с ExxonMobil, Европейской организацией по ядерным исследованиям (ЦЕРН) и Fermilab, для поиска максимального числа задач, в которых применение подобных интегрированных квантовых систем было бы эффективно.

Такое сотрудничество в сочетании с прогрессом в технологии квантовых вычислений способствует росту этой зарождающейся индустрии. Так, по прогнозам консалтинговой компании MarketsandMarkets, глобальный рынок квантовых вычислений увеличится с 93 млн долл. в 2019 г. до 283 млн долл. в 2024. За указанный период рынок QaaS (квантовые вычисления, как сервис), охватывающий приложения, связанные с задачами оптимизации, искусственного интеллекта и исследованием материалов, вырастет с 4 млн до 13 млн долл. По прогнозам, наиболее быстрыми темпами будут расти сегменты рынка, связанные с консалтинговыми услугами для космических исследований, обороны (для создания надежной высокоскоростной связи) и исследования материалов [40].

5.7. Заключение

Повышенный интерес к квантовым вычислениям, наблюдаемый в последние годы, обусловлен разработкой ряда эффективных алгоритмов квантовых вычислений, теоретически позволяющих решить ряд важных задач, для которых не существует эффективных алгоритмов решения на традиционных компьютерах.

Наиболее продвинутое направление исследований в области создания квантовых компьютеров: связано с использованием сверхпроводниковых кубитов. В последние годы расширяется использования этого класса компьютеров для оценки возможности решения практических задач в наукоемких компаниях и отраслях.

Наметилась гонка за лидерство в области квантовых технологий, главным образом, в области квантовых коммуникаций, в которой участвуют ведущие экономики мира.

Для создания полноценного квантового компьютера в обозримом будущем необходимо решение сложных принципиальных проблем, связанных с преодоления эффектов декогерентизации и ошибок в схемах.

Список литературы к главе 5

1. Варшавский Л.Е. Текущие проблемы развития квантовых вычислений // Теория и практика институциональных преобразований в России: сб. научных трудов / Ерзнкян Б.Г. (ред.). Вып. 47. М.: ЦЭМИ РАН, 2019. С. 33–45.
2. Валиев К. А. Квантовые компьютеры и квантовые вычисления // Успехи физических наук. 2005. Т. 175. № 1. С. 1–39.
3. URL: <http://www.carolla.com/quantum/QuantumComputers.htm>
4. Quantum computers. URL: http://ewh.ieee.org/r10/bombay/news4/Quantum_Computers.htm
5. Перри Р. Элементарное введение в квантовые вычисления. Долгопрудный: ИД «Интеллект», 2015.
6. Нильсен М., Чанг И. Квантовые вычисления и квантовая информация. М.: Мир, 2006.
7. Килин С.Я. Квантовая информация // Успехи физических наук. 1999. Т. 169. № 5. С. 508–526.
8. Wolf E.L. Quantum Nanoelectronics: An Introduction to Electronic Nanotechnology and Quantum Computing. Wiley, 2009.
9. URL: <https://spectrum.ieee.org/computing/hardware/google-plans-to-demonstrate-the-supremacy-of-quantum-computing>
10. URL: <https://habr.com/ru/post/435560/>
11. Rachel Courtland A First for Quantum Computing? URL: <http://spectrum.ieee.org/tech-talk/computing/hardware/a-first-for-quantum-computing>
12. Hardy Q. A Strange Computer Promises Great Speed. URL: http://www.nytimes.com/2013/03/22/technology/testing-a-new-class-of-speedy-computer.html?_r=0&page-wanted=print
13. Choi Ch. Google and NASA Launch Quantum Computing AI Lab. May 16, 2013. URL: <http://www.technologyreview.com/news/514846/google-and-nasa-launch-quantum-computing-ai-lab/>
14. URL: <https://spectrum.ieee.org/tech-talk/computing/hardware/dwave-launches-free-quantum-cloud-service>
15. URL: <https://spectrum.ieee.org/cars-that-think/transportation/self-driving/ford-signs-up-to-use-nasas-quantum-computers>
16. URL: <https://eandt.theiet.org/content/articles/2019/04/how-quantum-computers-are-transforming-travel/>
17. URL: <https://quantumcomputingreport.com/news/chinese-scientists-set-new-quantum-factoring-record/>
18. Scientists create first electronic quantum processor. URL: <http://www.physorg.com/news165418586.html>
19. Three Solid-State Qubits Entangled: Big Step Toward Quantum Error Correction. URL: <http://www.sciencedaily.com/releases/2010/09/100929132533.htm>
20. Quantum Computing Research Edges Toward Practicality. URL: <http://www.sciencedaily.com/releases/2010/10/101005104446.htm>
21. URL: <http://www.nanowerk.com/news2/newsid=26603.php>
22. URL: <http://www.nanowerk.com/news/newsid=24827.php>
23. URL: <https://www.bcg.com/en-ru/publications/2018/next-decade-quantum-computing-how-play.aspx>
24. Maslov D., Nam Y., Kim J. An Outlook for Quantum Computing // Proceedings of the IEEE. 2019. Vol. 107. No. 1. Pp. 5–10.
25. URL: <https://spectrum.ieee.org/tech-talk/computing/networks/quantum-drone>

26. URL: <https://www.washingtonpost.com/business/2019/08/18/quantum-revolution-is-coming-chinese-scientists-are-forefront/?noredirect=on>
27. URL: <https://www.bcg.com/en-ru/publications/2018/next-decade-quantum-computing-how-play.aspx>
28. URL: <https://spectrum.ieee.org/tech-talk/computing/hardware/race-for-the-quantum-prize-rises-to-national-priority>
29. URL: <https://www.gov.uk/government/news/new-funding-puts-uk-at-the-forefront-of-cutting-edge-quantum-technologies>
30. URL: http://www.tadviser.ru/index.php/%D0%A1%D1%82%D0%B0%D1%82%D1%8C%D1%8F:%D0%9A%D0%B2%D0%B0%D0%BD%D1%82%D0%BE%D0%B2%D1%8B%D0%B9_%D0%BA%D0%BE%D0%BC%D0%BF%D1%8C%D1%8E%D1%82%D0%B5%D1%80_%D0%B8_%D0%BA%D0%B2%D0%B0%D0%BD%D1%82%D0%BE%D0%B2%D0%B0%D1%8F_%D1%81%D0%B2%D1%8F%D0%B7%D1%8C
31. URL: <https://www.ibm.com/blogs/research/2018/10/quantum-advantage-2/>
32. URL: <https://quantumcomputingreport.com/news/ibm-achieves-higher-quantum-volume-with-the-ibm-q-system-one-design/>
33. URL: <https://spectrum.ieee.org/computing/hardware/the-case-against-quantum-computing>
34. URL: <https://spectrum.ieee.org/tech-talk/computing/hardware/the-us-national-academies-reports-on-the-prospects-for-quantum-computing>
35. URL: <https://spectrum.ieee.org/tech-talk/computing/hardware/an-optimists-view-of-the-4-challenges-to-quantum-computing>
36. Pitt professor lays groundwork for quantum computers. URL: <http://www.post-gazette.com/pg/10284/1094223-115.stm>
37. URL: http://www.nytimes.com/2013/03/22/technology/testing-a-new-class-of-speedy-computer.html?_r=0&pagewanted=print
38. Simonite T. The CIA and Jeff Bezos Bet on Quantum Computing. URL: <http://www.technologyreview.com/news/429429/the-cia-and-jeff-bezos-bet-on-quantum-mputing/>
39. URL: <https://spectrum.ieee.org/view-from-the-valley/at-work/start-ups/is-it-time-for-quantum-computing-startups>
40. URL: <https://www.marketsandmarkets.com/Market-Reports/quantum-computing-market-144888301.html>

РАЗДЕЛ 3.

АНАЛИЗ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ПРОБЛЕМ, СВЯЗАННЫХ С ИКТ

ГЛАВА 6.

ИССЛЕДОВАНИЕ ФАКТОРОВ РИСКА ИКТ НА УРОВНЕ ЭКОНОМИЧЕСКИХ АГЕНТОВ

6.1. Введение

Как уже отмечалось в данной работе, побочным эффектом форсированного распространения этих технологий стало появление все новых проблем экономического, правового, этического характера и связанных с ними киберугроз и рисков, которые в перспективе начнут сдерживать внедрение ИКТ. Учитывая уроки начала 2000-х гг., когда молниеносно подхватывались непроверенные предложения и инструменты финансовых «инноваторов», приведшие к финансовому кризису 2008–2009 гг., необходим тщательный анализ социально-экономических последствий решений и предложений, рекомендуемых специалистами в области ИКТ и поощряющими их инвесторами, не заботящимися о долгосрочной перспективе [1].

Эти проблемы весьма остро стоят и в нашей стране, тем более что в 2017 г. была принята Программа «Цифровая экономика Российской Федерации». Цель настоящей главы состоит в анализе экономических проблем, связанных с расширением использования ИКТ, учет и решение которых могли бы способствовать смягчению неблагоприятных для развития экономических агентов последствий. Значительное внимание уделяется анализу факторов риска, связанных с развитием производственной базы ИКТ-компаний и предприятий, производящих аппаратные средства ИКТ. Рассматриваются экономические факторы риска использования ИКТ для экономических объектов, а также игровой подход к исследованию динамики показателей рынков, участники которых подвергаются кибератакам на производственную инфраструктуру. Материал данной главы основан на статьях автора [2, 3].

6.2. Экономический анализ факторов риска производителей аппаратных средств ИКТ

К настоящему времени накоплен определенный опыт исследования хозяйственных рисков в рыночных условиях. Так, в монографии [4] дана классификация хозяйственных факторов риска (внешних и внутренних) производственного предприятия, а также предложен комплекс мер по управлению (менеджменту) рисками. В настоящей работе проведен анализ основных факторов рисков, которые учитывают в своей деятельности ведущие компании в области ИКТ (Intel, AMD, Cisco, Apple, IBM, TSMC). Такой анализ, проведенный на основе информации, содержащейся в годовых отчетах упомянутых компаний за период с начала 2000-х по 2014 г., позволил выделить следующие группы рисков:

- производственные;
- технологические;
- рыночные;
- нерыночные;
- фискальные;
- финансовые;
- медицинские и экологические;
- криминальные;
- геополитические.

6.2.1. Производственные риски

К производственным рискам, к которым наиболее чувствительны ведущие компании ИКТ, можно отнести:

- нарушения в цепи поставок сырья, материалов, компонентов и оборудования;
- нарушения в цепи распределения и реализации продукции компании;
- отклонение значений технико-экономических показателей и показателей качества продукции от ожидаемых;
- профессиональные качества и мотивация персонала и др.

Практически все исследуемые компании указывают на наличие рисков нарушения в цепи поставок сырья, материалов, компонентов и оборудования, а также в цепи распределения и реализации продукции. Большая обеспокоенность возможностью подобных нарушений в компаниях микроэлектронного профиля (Intel, TSMC) обусловлена зависимостью этих производителей от сотен и даже тысяч поставщиков

сырья, материалов, компонентов и оборудования, многие из которых, в условиях глобализации, разбросаны по всему миру.

С другой стороны, поставки некоторых видов сырья на предприятия ряда компаний (например, TSMC) монополизированы одним производителем. Большая часть производства микропроцессоров и графических процессоров, разрабатываемых компанией AMD, после 2009 г. сосредоточена на фабриках компании Global Foundries (GF). Число потребителей готовой продукции также в ряде компаний невелико (так, на 10 потребителей микросхем компании TSMC приходится более 60% всего ее оборота, причем на крупнейшего из них – приблизительно 20%).

На риски, связанные с зависимостью от компаний-разработчиков software, указывают также компании Cisco и Apple. Кроме того, Cisco указывает на зависимость роста своих доходов от развития Интернета и связанных с ним систем. Следует отметить и зависимость показателей деятельности компаний-разработчиков микропроцессоров (AMD, Intel) от уровня деловой активности компаний-смежников, производящих материнские платы, устройства памяти, программное обеспечение (software) для компьютеров. Кроме того, компании AMD, Cisco, Apple отмечают зависимость своего производства от использования лицензий и интеллектуальной собственности третьих сторон. AMD указывает на решающую роль компании Microsoft, как разработчика software, в поддержке своей продукции, лишившись такой поддержки компания столкнется с существенными материальными потерями.

Все компании отмечают важность привлечения на работу профессиональных сотрудников, что объясняется высокими науко- и капиталоемкостью производимой ими продукции.

6.2.2. Технологические риски

К факторам, связанным с технологическими рисками ИКТ-компаний, нами, в частности, отнесены следующие:

- неудачи в разработке и запуске процессных и продуктовых технологий;
- появление новых передовых технологий у конкурентов;
- зависимость от интеллектуальной собственности других компаний.

Практически все исследуемые компании указывают на решающую роль новых, востребованных потребителями продуктов и технологий для успешного ведения своего бизнеса. При этом Intel предупреждает, что не все разрабатываемые его учеными продукты и технологии могут оказаться коммерчески успешными.

6.2.3. Рыночные риски

К факторам, связанным с этими рисками, отнесены, в частности:

- острая конкуренция;
- отсутствие спроса на продукцию;
- превышение предложения над спросом;
- реализация ошибочной стратегии развития и др.

Вышеуказанные факторы универсальны и оказывают влияние на деятельность всех компаний, функционирующих в рамках рыночной системы, независимо от отрасли.

6.2.4. Нерыночные риски

К факторам, связанным с этими рисками, в работе отнесены, в частности:

- поддержка конкурентов государством или третьими сторонами;
- возникновение судебных тяжб;
- отказ потенциальных государственных заказчиков от будущих контрактов или заказов;
- внедрение регуляторами рынка невыгодных для компаний тарифов и др.

Так, Intel указывает на свою частичную зависимость от государственных контрактов, величина которых может быть подвержена сокращениям из-за трудно предсказуемых возникающих бюджетных ограничений. Кроме того, из-за возможности несоблюдения требований, оговоренных в договорах на выполнение контрактных работ, относящихся, в частности, к ведению бухгалтерского учета, режиму секретности, к компании могут быть применены санкции (материальные, судебного характера), вплоть до отмены контракта, что не сможет не отразиться на экономических результатах ее деятельности.

Все исследуемые компании отмечают риски возникновения негативных последствий разного рода судебных тяжб, связанных, главным образом, с обвинениями в нарушении прав на интеллектуальную собственность.

6.2.5. Фискальные риски

Все компании отмечают такие факторы этих рисков, как частое изменение налоговой системы; установление невыгодных ставок налогов и др.

6.2.6. Финансовые риски

Компании Cisco, Apple, TSMC указывают на финансовые риски, связанные, главным образом, со следующими событиями:

- неопределенностью результатов финансового инвестирования;
- высокой волатильностью фондовых индексов (динамика этих индексов за период 2004–2014 гг. представлена на рис. 1);
- высокой волатильностью обменных курсов;
- неблагоприятными ставками кредитов и др.

Так, компании Intel и TSMC осуществляют стратегические инвестиции, которые могут оказаться неприбыльными. В частности, эти компании вложили значительную часть инвестиций в облигации компании ASML, длительное время разрабатывавшей установки EUV фотолитографии, с которыми связывались перспективы перехода на производство микросхем с топологическими размерами 7 нм и менее.

Высокая волатильность фондовых индексов, особенно индивидуальных фондовых индексов компаний, затрудняет разработку и реализацию стратегий их развития. Она проявляется в слабой связи, или отсутствии таковой, между показателями эффективности производства (в частности, операционной или чистой прибылью) и фондовыми индексами компаний.

Проведенный нами анализ показал, незначимую или слабую статистическую связь между этими показателями для всех исследованных компаний, за исключением компании IBM. Для этой компании получена следующая зависимость между ее фондовым индексом S_{IBM_t} , величиной чистой прибыли в млрд долл. NI_t , а также индексом $S \& P_{IT_t}$ для ИТ-компаний:

$$S_{IBM_t} = 19,192 NI_t + 0,827 S \& P_{IT_t} - 88,151, \quad R^2 = 0,913, \quad (6.1)$$

(3,331) (0,202) (39,642)

где t обозначает год (в скобках под оценками параметров приведены оценки их среднеквадратичных отклонений).

В то же время проведенный анализ показывает, что фондовый индекс компании Intel S_{Intelt} значимо связан лишь с технологическим индексом Доу-Джонса DJ_{techt} :

$$S_{Intelt} = 0,530 DJ_{techt} + 23,967, \quad R^2 = 0,712, \quad (6.2)$$

(0,097) (12,226)

Более того, значение этого индекса в определенной степени влияет на динамику капитальных вложений компании. Так, эконометрический анализ показал наличие следующей связи между объемом капитальных вложений $Capex_{Intelt}$ в млн долл. и технологическим индексом Доу-Джонса DJ_{techt} , а также величиной операционной прибыли OI_t (также в млн долл):

$$Capex_{Intel_t} = \frac{28,984}{(z-0,171)^2} DJ_{tech_t} + \frac{0,450}{(z-0,174)^2} OI_t, \quad (6.3)$$

где z – оператор сдвига, т.е. $zx_t = x_{t+1}$.

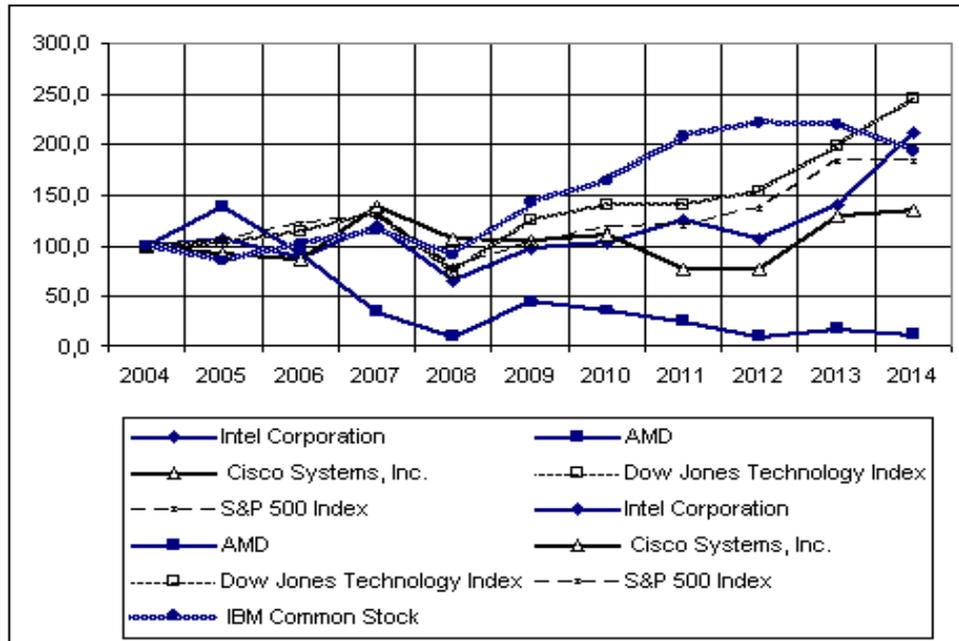


Рис. 6.1. Динамика фондовых индексов компаний ИКТ, а также индексов Доу-Джонса и S&P500 (составлено по данным годовых отчетов компаний)

Об удовлетворительном качестве полученной зависимости свидетельствует приемлемое значение коэффициента вариации ($v = 0,100$, см. также рис. 6.2). Это позволяет предположить, что компания Intel при планировании своей деятельности в большей степени ориентируется на технологический индекс Доу-Джонса DJ_{tech_t} , чем на другие индексы.

В то же время другая в прошлом капиталоемкая компания, AMD, при осуществлении капитальных вложений $Capex_{AMD_t}$ с конца 1990-х по 2009 г., повидимому, руководствовалась значениями собственного фондового индекса $S_{AMD_{t-1}}$, о чем свидетельствует достаточно устойчивая статистическая зависимость между этими показателями:

$$Capex_{AMD_t} = 0,319 Capex_{AMD_{t-1}} + 7,515 S_{AMD_{t-1}}, \quad R^2 = 0,911. \quad (6.4)$$

(0,255) (2,898)

Для остальных исследовавшихся ИКТ-компаний, являющихся существенно менее капиталоемкими, такой связи не обнаружено.



Рис. 6.2. Динамика фактических и расчетных (Оценка) на основе зависимости (3) объемов капитальных вложений в компании Intel (млн ед.)

Таким образом, в капиталоемких ИКТ-компаниях фондовые индексы оказывают влияние на динамику капитальных вложений и производства. В связи с этим, учет финансовых факторов риска, связанных с волатильностью фондовых индексов в таких компаниях требует особого внимания.

6.2.7. Экологические риски

Все компании выделяют следующие виды таких рисков, связанные, главным образом, со следующими действиями регулирующих органов в области экологии:

- запретом на использование опасных для окружающей среды и здоровья людей реагентов, материалов и технологических процессов;
- установлением более жестких норм на выбросы в окружающую среду вредных веществ, а также на источники воздействия на окружающую среду (разного рода излучения) и др.

В частности, Intel указывает на риски, связанные с регулированием природоохранными органами использования новых реагентов и материалов, применяемых компанией при производстве своих микросхем. Кроме того, ведущего производителя микропроцессоров беспокоит возможность ужесточения или даже запрета регулируемыми органами одного из используемых им газов – perfluorocompounds, сокр. PFCs (это событие может произойти в случае принятия в США программы мероприятий, направленных на уменьшение воздействия производства на изменения климата).

6.2.8. Криминальные риски

Все компании отмечают риски, связанные с повышением интенсивности кибератак, взломом данных нарушением работы информационных систем, а также нарушением прав компаний на интеллектуальную собственность. Эти действия злоумышленников могут испортить репутацию компаний, а для устранения последствий от хакерства и промышленного шпионажа могут потребоваться значительные материальные ресурсы, что в итоге может осложнить положение компании.

Вместе с тем AMD и Cisco отмечают риски, связанные с появлением на рынке контрафактной версии на свои продукты.

6.2.9. Геополитические риски

Все компании выделяют риски, связанные со следующими событиями:

- военными действиями в зоне производства;
- введением санкций на поставку критически важного сырья, материалов и оборудования;
- отказом в выдаче лицензии на производство или использование необходимого ресурса и др.

6.3. Анализ факторов риска потребителей ИКТ-продуктов и услуг

6.3.1. Недостаточно проработанные решения разработчиков ИКТ

В связи с приближением топологических размеров микросхем к предельно допустимым уровням, возрастают **и технологические риски** компаний-производителей элементной базы. Так, даже ведущий производитель Intel предупреждает, что не все разрабатываемые его учеными продукты и технологии могут оказаться коммерчески успешными.

Следует отметить, что некоторые недоработки продукции производителей становятся известными через много лет после начала эксплуатации техники. Так, в самом конце 2017 г., после долгих лет эксплуатации были обнаружены уязвимости микропроцессоров, получившие название Meltdown (поражает микропроцессоры Intel) а также Spectre (поражает микропроцессоры PC и мобильных устройств).

В этой связи, нельзя не согласиться с мнением специалиста по кибербезопасности С. Боргом (S. Borg), который отметил, что инженеры-разработчики «железа» (hardware) должны оценивать свои создаваемые продукты с позиции **максимизации**

затрат у возможных хакеров^{ix}. Отметим, что подобный подход к проектированию изделий известен в ряде других областей науки и техники, например, в атомной энергетике.

6.3.2. Экономические потери, связанные с ИКТ

Несмотря на бесспорно огромную общественную полезность информатизации, следует отметить, что часто внедрение информационно-коммуникационных технологий рассматривается как самоцель, причем не производится оценки эффективности таких мероприятий. В результате потребители информационных продуктов и услуг стали «заложниками коммерческих интересов их производителей и продавцов и наша от них зависимость гораздо сильнее нашего на них влияния». По мнению ряда аналитиков, нынешний этап информатизации отличается односторонним характером сложившихся отношений. Способствует этому, в частности, несовместимость ИКТ-продукции разных конкурентов, малый жизненный цикл изделий (как hardware, так и software), а также агрессивная маркетинговая политика. В силу последних факторов, потребителям приходится часто обновлять свои продукты, следствием чего являются неоправданные во многих случаях затраты ресурсов и труда (в значительной мере, на переобучение) [5, с. 78–79.].

Так, в [5] в качестве примера приводится федеральный проект создания универсальных электронных карт граждан России, которые должны были обеспечить гражданам страны универсальный доступ к государственным услугам, а в дальнейшем могли бы быть использованы вместо удостоверений личности и водительских прав. Однако указывается, что при рассмотрении и принятии этого проекта не был проведен должным образом его аудит. В итоге бюджет на его выполнение возрос с первоначальной оценки в 150–160 млрд до 450 млрд руб. [5, с. 176–178.].

Неправильный учет затрат на государственные ИТ-проекты с большим объемом работ по созданию software характерен для всех стран. Так, в США первоначальные затраты на стартовавшую в 2003 г. Программу боевых систем будущего (Future Combat Systems Program – FCC), в которой намечалось осуществить модернизацию вооруженных сил путем создания информационно-коммуникационных сетей, роботизированных и беспилотных средств, оценивались в 92 млрд долл. Впоследствии предполагаемая величина затрат на программу, вероятность успеха которой даже по мнению некоторых высокопоставленных заказчиков составляла 28%,

^{ix} URL: http://spectrum.ieee.org/view-from-the-valley/computing/embedded-systems/why-hardware-engineers-have-to-think-like-cybercriminals-and-why-engineers-are-easy-to-fool?utm_source=Tech+Alert&utm_medium=Email&utm_campaign=TechAlert_05-18-17&bt_ee=JwZmn+IWfgWSTn5HrwTe80uSLDZij+gK3as1tqUqPm67Aruejk2nYd65i6XNH9rn&bt_ts=1495116783244 (дата доступа: 23.05.2018).

повысилась до 200–300 млрд долл. [6]. Специалисты отмечают, что за последнее десятилетие практически не наблюдается позитивных изменений в оценке эффективности, планировании и управлении ИТ-проектами [7].

6.3.3. Экономические факторы риска для экономических организаций

В предыдущем пункте указывалось, каких последствий опасаются ИКТ-компании в случае взлома своих информационных систем. Такие опасения в равной степени разделяются всеми экономическими объектами. В связи с этим, возрастают расходы всех организаций на обеспечение надежности своей информационной системы и информационной безопасности. Целесообразная величина затрат на эти мероприятия зависит от величины стандартной меры риска, определяемой как произведение вероятности риска на значение ущерба. При этом суммарные затраты организации на обеспечение надежности своей информационной системы и информационной безопасности зависят от ее специфики и структуры выполняемых работ. В табл. 6.1 приведены типичные ИТ-риски в финансовых организациях (составлена на основе [8]).

В крупных финансовых и промышленных компаниях существенная часть операционных и капитальных затрат приходится на сооружение и эксплуатацию центров обработки данных (ЦОД), электрическая мощность которых может достигать десятков и даже сотен мегаватт (так, электрическая мощность ЦОД Supernap в Лас-Вегасе, занимающего общую площадь приблизительно 40 тыс. м², составляет 250 МВт), а капитальные затраты доходят до 500 млн долл. [9].

Таблица 6.1

Типичные ИТ-риски в финансовых организациях [2]

| Риск | Вероятность | Ущерб |
|--|--|------------------------------|
| Авария в центре обработки данных (ЦОД) | Зависит от надежности энергообеспечения, местоположения ЦОД (в зонах, подверженных стихийным бедствиям: землетрясениям, цунами, наводнениям и др.) | Значительный |
| Потеря важных данных в результате взлома | Зависит от принятых мер по информационной безопасности и степени важности данных | Значительный |
| Сетевые угрозы и уязвимости | Зависит от принятых мер по информационной безопасности | Значительный |
| Кража файлов, шпионаж | Зависит от принятых мер по информационной безопасности | Значительный |
| Операционный риск, связанный с ошибками персонала | Незначительная | От среднего до значительного |
| Операционный риск, связанный с вынужденным простоем информационной системы | Незначительная | От среднего до значительного |

В связи с необходимостью обеспечения надежности работы ЦОД, они классифицируются по 4 уровням надежности (Tier 1–4). Наиболее высоким является уровень надежности Tier 4. Так, Uptime Institute относит к этому уровню ЦОД, обеспечивающие за счет многократного дублирования (резервирования) всех критических элементов инфраструктуры (включая системы энергообеспечения и охлаждения) бесперебойную работу с показателем надежности (отказоустойчивости) 99.995% (при этом время простоя составляет 0.438 часа за год). Уровни Tier 3, Tier 2, Tier 1 характеризуются показателями надежности соответственно 99,982%, 99,741 и 99,671%. Им соответствуют времена простоя (часов за год): 1,577, 22,688 и 28,817. Естественно, что повышение надежности ЦОД за счет многократного резервирования ключевых блоков требует значительных дополнительных затрат.

Вместе с тем, растут затраты и на обеспечение ИТ-безопасности (IT security) или кибербезопасности. Подобный рост обусловлен интенсификацией киберпреступлений и повышением числа успешных кибератак. Этому во многом способствуют бурное и, вместе с тем, стихийное развитие облачных сервисов и Интернета вещей (Internet of Things – IoT). Переход предприятий и организаций на облачные вычисления позволит им, по мнению аналитиков, уменьшить затраты на ИТ-сектор (включая экономию электроэнергии), повысить «проворность» (англ. agility), приспособляемость к переменам. Следует, однако, отметить, что во многих случаях потенциал облачных сервисов в полной мере не используется из-за нерешенности проблем компьютерной безопасности. Даже операторы ЦОД признают, что нет абсолютных средств безопасности, которые могли бы спасти их от преднамеренной атаки [10].

Еще большей уязвимостью к кибератакам обладает Интернет вещей, который стал главным драйвером развития информационных технологий. Несмотря на то, что многие руководители компаний считают, что Интернет вещей поможет им повысить конкурентоспособность, переход на IoT сдерживается недостаточной решенностью проблем кибербезопасности, использования данных и совместимости приборов. В то же время, некоторые эксперты указывают и на отсутствие ясных выгод для компаний от такого перехода [11].

Определенную настороженность бизнеса вызывают и участившиеся в последние месяцы успешные кибератаки на компании, использующие IoT. Так, в январе 2016 г. была совершена кибератака на электрические сети Израиля, приведшая к отключению большей части компьютеров израильского департамента энергоснабжения. В конце 2015 г. хакерами была выведена из строя электросеть компании «Прикарпатьеоблэнерго» в результате чего значительная часть Прикарпатья и г. Ивано-Франковск (Украина) остались без электричества [12].

С помощью устройств, относящихся к Интернету вещей, были совершены DDoS-атаки на российские банки^x. При этом использовались бытовые приборы, подключенные к Интернету. Как отмечает руководитель исследовательского центра «Лаборатории Касперского» Ю. Наместников, «...единственный способ защитить устройство Интернета вещей – это сделать его безопасным с самого начала, by-design, то есть оно должно быть произведено с учетом требований безопасности. Поэтому вендорам необходимо уделять больше внимания безопасности умных устройств на этапе производства. Регуляторам же нужно создавать законодательную базу, новые стандарты для использования подобных устройств» [13].

В октябре 2016 г. подверглась мощнейшей DDoS-атаке со стороны ботнета с условным названием Mirai компания Дун, предоставляющая сервисные услуги компаниям с наиболее посещаемыми вебсайтами мира. В атаке были использованы десятки миллионов приборов, связанных с Интернетом вещей. В итоге для многих пользователей США, и даже Австралии, оказались недоступны такие популярные сайты, как Twitter, Spotify, Reddit and CNN [14].

Российские компании также являются объектами интенсивных кибератак хакеров. Так, компания «Россети» ежегодно подвергается 9 млн кибератакам, что составляет 30 тыс. в день. Проникновение в технологические системы управления осуществляется, в частности, через электронную почту, мобильные сервисы, портал подачи заявок^{xi}.

По данным Group-IB в 2019 г. российские организации подвергались примерно 50 тыс. кибератакам. В связи с повышением эффективности систем кибербезопасности финансовых организаций, увеличилось число хакерских атак на промышленные объекты и объекты критически важной инфраструктуры^{xii} с целью кражи денежных средств. «По данным «Лаборатории Касперского», в 2018 г. 48% компьютеров автоматизированных систем управления технологическим процессом (АСУ ТП) предприятий подверглись кибератакам с целью вымогательства

По некоторым экспертным оценкам, за один год потери российской экономики из-за кибератак сопоставимы с затратами по Программе «Цифровая экономика». Так, Сбербанк оценивал сумму потерь российской экономики от кибератак в 2019 г. в 1,6–1,8 трлн руб.^{xiii}. Успешному проведению кибератак способствует то,

^x DDoS-атака (distributed denial of service – англ. распределенный отказ в обслуживании) – атака, вызывающая перегрузку на сервере или сайте жертвы путем генерации большого числа входящих запросов со стороны хакерской сети, в которой объединено большое число компьютеров.

^{xi} URL: http://www.ng.ru/economics/2019-10-09/100_110709102019.html (дата доступа 10.09.2019).

^{xii} К критически важной инфраструктуре принято относить системы, сети, объекты и службы, нарушение работы которых непосредственно влияет на здоровье, безопасность и благосостояние граждан.

^{xiii} URL: http://www.ng.ru/economics/2019-11-05/1_7718_hackers.html (дата доступа 15.01.2020).

что многие промышленные объекты оснащены старыми АСУ ТП, при разработке которых не учитывалась возможность масштабных киберпреступлений^{xiv}.

Следует отметить, что особенно быстрый рост киберпреступлений фиксируется в последние годы. Так, еще в 2017 г. каждый третий гражданин США был информирован о взломе своих персональных данных (в 2016 г. – каждый восьмой).

В связи с интенсификацией хакерских атак и киберпреступлений, существенно возрастают расходы всех организаций на обеспечение надежности своей информационной системы и кибербезопасности. Так, в США среднегодовые затраты компаний численностью свыше 1000 сотрудников в 2015 г. только на борьбу с киберпреступлениями составляли 15 млн долл. Значительная величина затрат связана с тем, что в среднем требовалось 46 дней для устранения последствий успешной для преступников кибератаки [15].

При этом следует отметить важную тенденцию, связанную с увеличением темпов прироста затрат на обеспечение кибербезопасности. Так, по данным консалтинговой компании Gartner, в 2016 г. общемировые затраты на обеспечение ИТ безопасности составили 81 млрд долл, то в 2019 г. они достигли почти 124 млрд долл. [16]. Правительство США выделило в государственном бюджете на 2017 г. 19 млрд долл. на обеспечение кибербезопасности, что на 35% больше всех федеральных расходов по этой статье, осуществленных в 2016 г. [17]. В 2016 г. расходы Сбербанка на кибербезопасность составили около 1,5 млрд руб., что вдвое превышает уровень подобных расходов 2014 г. [18.].

В докладе о рынке кибербезопасности (The Cybersecurity Market Report), подготовленном в первом квартале текущего года компанией Cybersecurity Ventures, прогнозировалось, что в 2017–2021 гг. величина расходов на кибербезопасность в мире превысит 1 трлн долл, причем среднегодовые темпы прироста таких затрат увеличатся и составят 12–15%. По оценкам этой компании только в 2017 г. на обеспечение кибербезопасности должно было быть затрачено 120 млрд долл, что в 35 раз превышает уровень 2004 г. Несмотря на такие расходы, в компании прогнозировали, что киберпреступления приведут в 2021 г. к глобальным убыткам в размере 6 трлн долл.

По оценкам другой консалтинговой компании, BI Intelligence, на обеспечение кибербезопасности в 2015–2020 гг. должно было быть затрачено 655 млрд долл., причем 386 млрд долл. – на обеспечение безопасности: персональных компьютеров; 172 млрд долл. – приборов для Интернета вещей и 113 млрд долл. – мобильных устройств.

^{xiv} URL: <https://www.vedomosti.ru/technology/articles/2019/04/17/799417-kolichestvo-kiberatak> (дата доступа 15.11.2019).

Многие из вышеприведенных прогнозных оценок оказались достаточно точными.

Интересные оценки содержатся в отчете «Кибервыгоды и риски: количественные оценки и прогнозирование баланса» (Cyber Benefits and Risks: Quantitatively Understanding and Forecasting the Balance), опубликованном Центром исследования будущего Университета Денвера (University of Denver's Pardee Center for International Futures). Так, в отчете отмечалось, что в 2019 г. годовые выгоды (benefits), связанные с инвестициями в ИКТ, должны были быть меньше, чем затраты на обеспечение кибербезопасности. Вместе с тем, кумулятивный эффект от использования ИКТ за все предыдущие годы существенно превосходит расходы на обеспечение кибербезопасности. На основе разработанной в Центре системы прогнозирования будущего, исследователи подготовили сценарии экономической эффективности ИКТ. В соответствии с одним из них, при недорогих ИКТ и системах безопасности, глобальный эффект в виде повышения производительности составит в 2030 г. 190 трлн долл. В соответствии с другим сценарием, при нестабильном и небезопасном Интернете потери мировой экономики составят 90 трлн долл. В связи с возможностью таких потерь, исследователи призывают правительства, бизнес и другие слои общества к кооперации с целью отражения возможных киберугроз. Они отмечают, что возможности Интернета, как и любого другого ресурса, могут быть исчерпаны. Интернет требует опеки (stewardship) для обеспечения позитивного вклада в развитие общества [19].

Наряду с рассмотренными выше рисками и угрозами, следует отметить также и опасности, связанные с распространением контрафактной ИКТ-продукции, которые могут иметь катастрофические последствия для потребителей. Так, потери только полупроводниковой промышленности США, связанные с распространением микроэлектронных контрафактных компонентов в 2011 г. оценивались в 7,5 млрд долл. Такие компоненты, источником которых обычно являются свалки неиспользованных приборов, отслуживших нормативные сроки, обычно поставляются для дорогостоящих устройств, характеризующихся длительным жизненным циклом (предназначенных для военной техники, для медицины и для других областей применения, в которых требуется высокая надежность изделий). Факторами, способствующими поставкам контрафактных компонентов, являются длинная цепь поставок для производства конечного устройства и сложности в проведении полноценного тестирования [20]. Однако главный фактор связан с несоблюдением этических норм отдельными поставщиками.

6.4. Моделирование динамики экономических показателей рынков при наличии киберугроз (на примере олигополии)

В настоящее время прямые и косвенные затраты в мире, связанные с киберпреступлениями, приближаются к 1 трлн долл., а их доля в мировом ВВП – к 1% [21]. По некоторым оценкам потери только российской экономики от кибератак в 2019 г. могли составить 1,6–1,8 трлн руб. [22].

Особую тревогу вызывают участвовавшие киберпреступления, направленные против критически важной инфраструктуры. В связи с форсированным внедрением ИКТ, осуществляемым без должного анализа рисков и угроз, связанных с развитием этих наиболее демократичных видов технологий, повышается роль исследований в области кибербезопасности, и, в частности, экономических аспектов, связанных с обеспечением приемлемых уровней обеспечения безопасности критической инфраструктуры, производства, а также товарных рынков, особенно рынков высокотехнологичной продукции. Вместе с тем, ввиду отсутствия надежных статистических данных о числе кибератак и вызванных ими потерь, возникают естественные трудности при исследовании реальных товарных рынков. Поэтому предлагаемые методические подходы и модели приходится проверять на примере условных рынков.

В настоящем п. рассматривается игровой подход к исследованию динамики показателей характерных для высокотехнологических производств олигополистических рынков, участники которых подвергаются кибератакам на производственную инфраструктуру. Анализируются результаты расчетов показателей условного рынка в соответствии с 3 сценариями.

6.4.1. Модель динамики показателей олигополистических рынков в условиях кибератак

Проводимый ниже анализ основан на использовании агрегированной динамической модели рационального поведения участников олигополии в виде линейной динамической игры по Нэшу–Курно с квадратичным критерием, в которой участвуют N фирм-олигополистов. Предполагается, что целью кибератак является уничтожение производственных мощностей участников рынка.

Центральным блоком модели является следующая зависимость, связывающая объемы товарного производства Q_{it} со входной переменной u_{it} (вводом мощностей), i – индекс фирмы, $i = 1, 2, \dots, N$, t – индекс года:

$$Q_{it} = W_i(z) u_{it} + Q_{0it} - \chi_i W_{0i}(z) v_{it}, \quad (6.5)$$

где $W_i(z) = B_i(z)/A_i(z)$ – передаточная функция, причем $A_i(z)$, $B_i(z)$ – полиномы относительно переменной z , представляющей собой оператор сдвига: $zx_t = x_{t+1}$:

$$A_i(z) = \sum_{k=0}^n a_{ik} z^k, \quad B_i(z) = \sum_{j=0}^m b_{ij} z^j, \quad m \leq n, \quad (6.6)$$

Q_{0it} – слагаемое, характеризующее начальные условия; $\chi_i = p(\mu_i)g_i$ – средняя величина падения производства в результате одной кибератаки; $p(\mu_i)$ – вероятность успешного отражения кибератаки, зависящая от μ_i – доли затрат в ОРЕХ на кибербезопасность; g_i – потери мощности и продукции в результате одной кибератаки; v_{it} – число кибератак; $W_{0i}(z)$ – передаточная функция, связывающая число кибератак с падением мощности и производства продукции. Другой блок модели – обратная функция спроса, связывающая цену на рынке P_t с суммарным предложением

$$Q_t = \sum_{i=1}^N Q_{it} : \quad P_t = a - bD_t = a - bQ_t, \quad (6.7)$$

где Q_{Ft} – суммарный объем производства малых компаний-ценополучателей; a , b – параметры.

Предполагается, что олигополисты используют скользящее планирование и в каждый момент времени τ максимизируют чистую текущую стоимость (NPV) с учетом того, что участники кибератак стремятся нанести компаниям максимальный ущерб:

$$J_{\tau i} = \sum_{t=\tau}^{\tau+T_p} \beta^t \left[(P_t - (1 + \mu_i)c_i)Q_{it} - q_i u_{it} - \frac{1}{2} \rho_{1i} u_{it}^2 + \frac{1}{2} \rho_{2i} v_{it}^2 \right] \rightarrow \max_{u_{it}} \min_{v_{it}}, \quad (6.8)$$

где $\beta = 1/(1+r)$ – дисконтирующий множитель, соответствующий ставке дисконтирования r ; P_t – цена продукции; c_i – средние производственные издержки (без амортизации); μ_i – соотношение между дополнительными затратами на кибербезопасность и средними производственными издержками; q_i – стоимость единицы мощностей; $\frac{1}{2} \rho_{1i} u_{it}^2$, $\frac{1}{2} \rho_{2i} v_{it}^2$ – затраты регулирования, характеризующие соответственно инвестиционные возможности олигополистов (см., например, [23]) и их восприятие интенсивности кибератак, с коэффициентами $\rho_{1i} > 0$, $\rho_{2i} > 0$, $i = 1, 2, \dots, N$; T_p – период скользящего планирования (для упрощения записи формул ставки налогов приняты равными нулю). Управляющими переменными для олигополистов в модели являются объемы ввода мощностей u_{it} , а также доли затрат на кибербезопасность μ_i , $i = 1, 2, \dots, N$.

В данном п. при проведении прогнозных расчетов использован подход к расчету оптимальных по Нэшу–Курно разомкнутых (open-loop), стратегий, основанный на представлении модели (6.5)–(6.8) в пространстве состояний и использовании обобщенных (generalized) матричных уравнений Риккати (см., гл. 2 и, например, [24, 25]). При этом модель (6.5)–(6.8) предварительно представлена в эквивалентной форме в пространстве состояний:

$$X_t = AX_{t-1} + B_i u_{it} + D_i v_{it}. \quad (6.9)$$

Используя подход к решению данной задачи, основанный на применении операционного исчисления (см. гл. 2, [26]), можно показать, что при равновесии по Нэшу–Курно для задачи (6.5)–(6.8) справедливы следующие соотношения (для упрощения формул, далее принято, что $Q_{0it} \equiv 0$, $i = 1, 2, \dots, N$):

$$v_{it} = \frac{\rho_{1i}}{\rho_{2i}} \chi_i \left[\frac{W_{0i}((\beta z)^{-1})}{W_i((\beta z)^{-1})} \right] u_{it}; \quad (6.10)$$

$$Q_{it} = \frac{\Gamma_i(z, (\beta z)^{-1})}{b} (p_t - PL_i), \quad (6.11)$$

где $PL_i = (1 + \mu_i)c_i + \frac{q_i}{W(1+r_i)}$ – приведенные затраты i -й фирмы;

$$\Gamma_i[z, (\beta z)^{-1}] = \frac{b \left[W_i(z)W_i((\beta z)^{-1}) - \frac{\rho_{1i}}{\rho_{2i}} \chi_i^2 W_{0i}(z)W_{0i}((\beta z)^{-1}) \right]}{\rho_{1i} + \left[W_i(z)W_i((\beta z)^{-1}) - \frac{\rho_{1i}}{\rho_{2i}} \chi_i^2 W_{0i}(z)W_{0i}((\beta z)^{-1}) \right]},$$

$$i = 1, 2, \dots, N; \quad (6.12)$$

$$P_t = \frac{1}{1 + \sum_{i=1}^N \Gamma_i[z, (\beta z)^{-1}]} \left\{ a + \sum_{i=1}^N \Gamma_i[z, (\beta z)^{-1}] PL_i \right\}.. \quad (6.13)$$

Из (6.10) следует, что чем больше величина $\frac{\rho_{1i}}{\rho_{2i}} \chi_i$, тем сильнее предполагаемое i -й компанией количество кибератак v_{it} связано со входной переменной u_{it} и соответственно, с производственными инвестициями ($Inv_{it} = q_i u_{it}$). Поэтому целесообразно принять эту величину в качестве индекса восприятия интенсивности кибератак $\left(CAAI_i = \frac{\rho_{1i}}{\rho_{2i}} \chi_i \right)$.

6.4.2. Результаты расчетов показателей условного олигополистического рынка

Предположения олигополистов об изменении индекса $CAAI_i$ могут быть положены в основу формирования сценариев развития рынков. На условном примере триополии ниже рассматриваются три сценария восприятия интенсивности киберугроз олигополистами: 1) оптимистический, при котором олигополисты руководствуются гипотезой о том, что с уменьшением вероятности успешных атак $p(\mu_i)$ (за счет роста доли затрат на кибербезопасность μ_i) относительная интенсивность атак, характеризуемая индексами $CAAI_i$, будет уменьшаться; 2) нейтральный, когда олигополисты при разработке стратегий исходят из гипотезы о стабильной относительной интенсивности атак (при этом индекс $CAAI_i = \frac{\rho_{1i}}{\rho_{2i}} \chi_i$ остается постоянным при уменьшении $p(\mu_i)$ и соответственно росте μ_i); 3) пессимистический, когда участники рынка руководствуются гипотезой о стабильной относительной динамике потерь, несмотря на снижение вероятности успешных атак $p(\mu_i)$ (при этом величина $\frac{\rho_{1i}}{\rho_{2i}} \chi_i^2 = CAAI_i \cdot \chi_i$ остается постоянной).

В качестве центрального блока модели рассматривается, модель освоения мощностей [27], для которой соотношения типа (6.5) для каждого олигополиста в пространстве состояний могут быть представлены в виде: (6.9), где

$$X_{it} = (x_{i1t}, x_{i2t}, x_{i3t})' - \text{вектор-столбец};$$

$$A_i = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0,95 \end{pmatrix}; \quad B_i = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}; \quad D_i = \begin{pmatrix} 0 \\ 0 \\ -\chi_i \end{pmatrix}; \quad (6.14)$$

$$Q_{it} = (k_0, k_1, 1)X_{it}, \quad 0 < k_0 < k_1 < 1, \quad i = 1, 2, \dots, N.$$

В базовом варианте для всех сценариев значения удельных производственных затрат на единицу производимой продукции (ОРЕХ) составляют $c_1 = c_2 = 100$, $c_3 = 85$; удельных капитальных вложений на единицу вводимой мощности $q_1 = q_2 = q_3 = 100$, доли затрат на кибербезопасность от величины ОРЕХ $\mu_1 = \mu_2 = \mu_3 = 0,05$, а процентной ставки – $r = 0,03$. Приняты следующие значения коэффициентов освоения мощностей: $k_0 = 0,4$; $k_1 = 0,7$. Значения коэффициентов ρ_{1i} , ρ_{2i} в базовом варианте представлены в табл. 6.2.

Таблица 6.2

Значения коэффициентов ρ_{1i}, ρ_{2i}

| i | ρ_{1i} | ρ_{2i} |
|-----|-------------|-------------|
| 1 | 70 | 4,3 |
| 2 | 70 | 4,3 |
| 3 | 50 | 3 |

Таким образом, третья компания, имеющая лучшие экономические показатели (c_3, ρ_{13}), является компанией-лидером. Принято также, что вероятность успешного отражения кибератак связана с долей затрат на кибербезопасность от величины ОПЕХ μ_i зависимостью $p_i = \exp(-40\mu_i)$, $i = 1, 2, 3$, а также, что во всех вариантах компании руководствуются одинаковой долей затрат на кибербезопасность $\mu_1 = \mu_2 = \mu_3$.

Расчеты показывают, что при первом и втором сценариях увеличение до определенного предела затрат компаний на кибербезопасность (в долях от операционных затрат) приводит к росту всех ключевых экономических показателей, что отражается на увеличении показателей чистой текущей стоимости (NPV) и отдачи от капитальных вложений (ROI , return on investment) во всех компаниях-участниках рынка. Дальнейший рост этих затрат приводит к снижению объемов производственных инвестиций и производства, а также показателей эффективности NPV и ROI . Вместе с тем, рыночная доля компаний-лидеров, имеющих меньшие удельные затраты на производство (Опех), может возрастать и с дальнейшим ростом доли выделяемых затрат на кибербезопасность. Динамика показателей триополии в первом сценарии представлена на рис. 6.3–6.7.

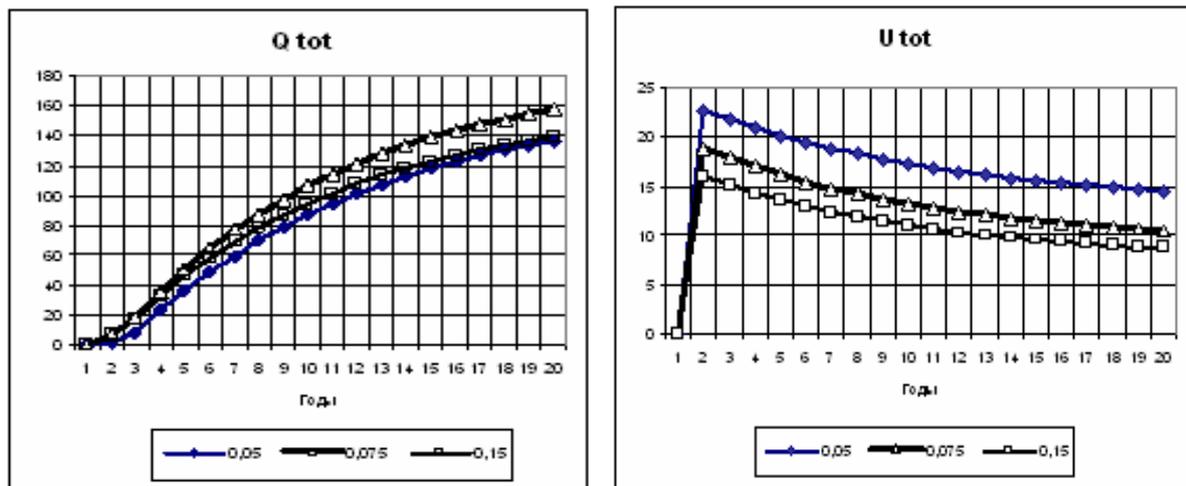


Рис. 6.3–6.4. Динамика суммарных объемов товарной продукции (Q_{tot}) и ввода мощностей (U_{tot}) в триополии при разных долях затрат на кибербезопасность в первом сценарии

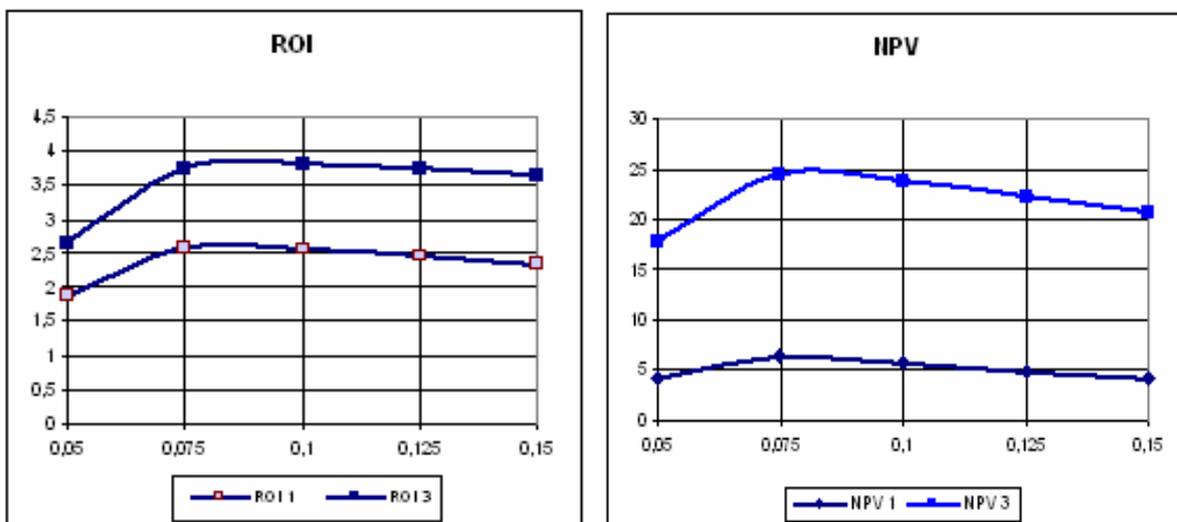


Рис. 6.5–6.6. Зависимость показателей эффективности за 20 лет (*ROI* и *NPV*) первой и третьей компаний в триополии от долей затрат μ на кибербезопасность в первом сценарии

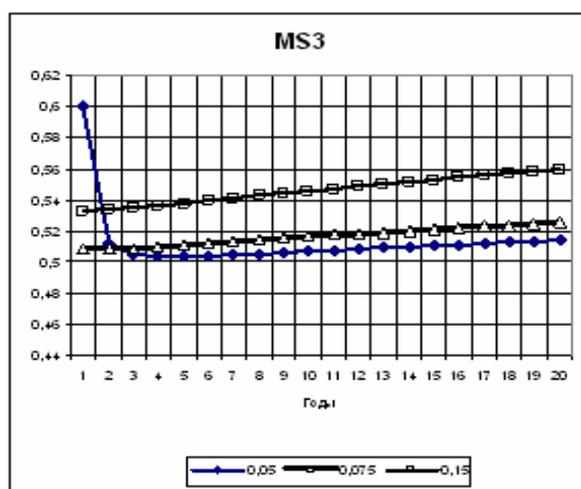


Рис. 6.7. Динамика рыночной доли третьей компании (MS 3) в триополии при разных долях затрат на кибербезопасность μ в первом сценарии

Во втором сценарии динамика показателей триополии имеет такой же характер, как и в первом сценарии. Вместе с тем, максимум показателей эффективности за 20-летний период в этом сценарии достигается при большей доле затрат μ на кибербезопасность (рис. 6.8).

В третьем сценарии одновременное повышение доли затрат на кибербезопасность в компаниях приводит к устойчивому уменьшению суммарных объемов инвестиций, производства (отгрузки), их прибыли, а также показателей *NPV* и *ROI* (рис. 6.9–6.12). Вместе с тем, происходит стабилизация рыночной доли компании-лидера.

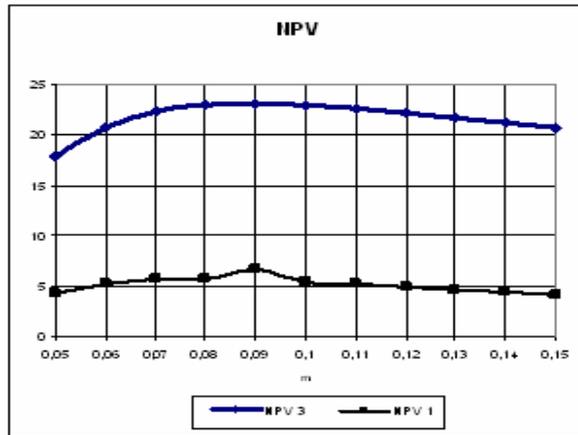


Рис. 6.8. Зависимость NPV за 20 лет первой и третьей компаний в триополии от долей затрат на кибербезопасность во втором сценарии

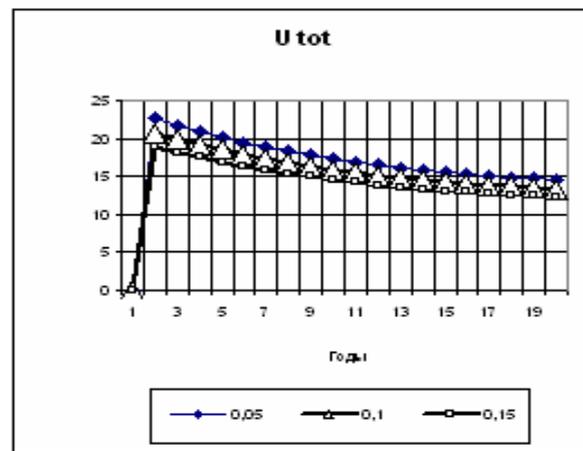
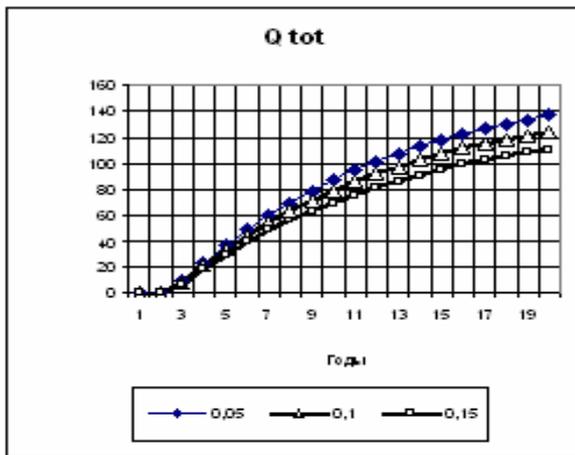


Рис. 6.9–6.10. Динамика суммарных объемов товарной продукции (Q_{tot}) и ввода мощностей (U_{tot}) в триополии при разных долях затрат на кибербезопасность в третьем сценарии

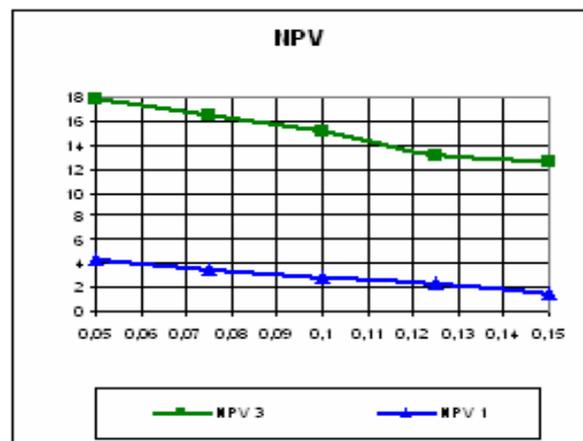
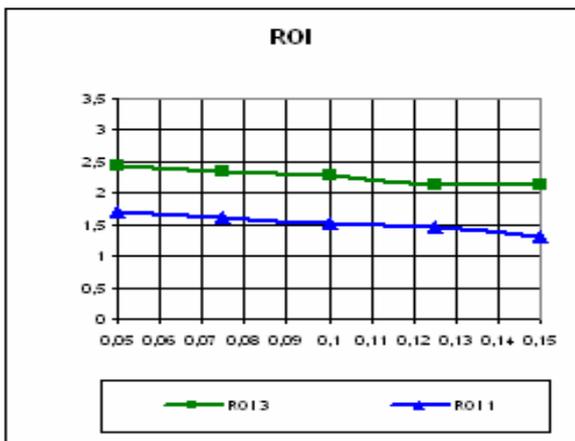


Рис. 6.11–6.12. Зависимость показателей эффективности за 20 лет (ROI и NPV) первой и третьей компаний в триополии от долей затрат μ на кибербезопасность в третьем сценарии

Следует отметить существенное отличие показателей эффективности в первых двух сценариях от показателей третьего сценария. Так, наибольшие значения NPV за 20-летний период для компаний в первом и втором сценариях достигаются при $\mu = 0,075$ и $\mu = 0,09$ и составляют соответственно 24,5 и 23,0 условных единиц. Этот же показатель в третьем сценарии составляет всего 17,8 условных единиц и соответствует базовому варианту (табл. 6.3).

Таблица 6.3

| Сценарий | $NPV1$ | $NPV3$ | μ |
|----------|--------|--------|-------|
| 1 | 6,368 | 24,501 | 0,075 |
| 2 | 6,649 | 23,032 | 0,09 |
| 3 | 4,295 | 17,786 | 0,05 |

Показатели в рассмотренных сценариях разительно отличаются от показателей в случае отсутствия кибератак. Так, как показывают расчеты, в этом случае для достижения тех же объемов производства Q_{tot} , что и в базовом варианте, потребовалось бы на 57% меньше производственных инвестиций (см. также рис. 6.13)!

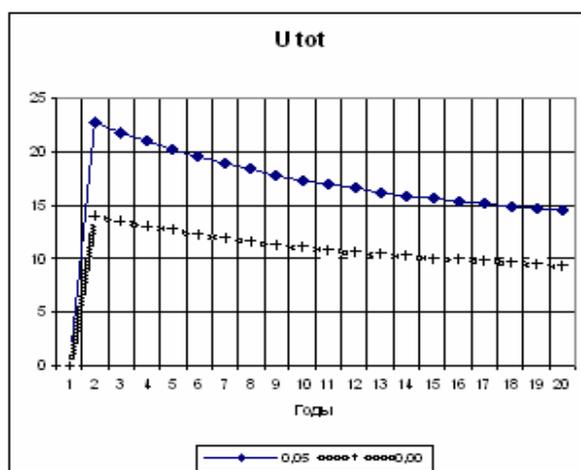


Рис. 6.13. Динамика суммарных объемов ввода мощностей (U_{tot}) в триополии в базовом варианте ($\mu = 0,05$) и в варианте отсутствия кибератак ($\mu = 0,00$)

Вместе с тем, при отсутствии кибератак объемы производства, соответствующие наиболее эффективному по критерию NPV варианту в первом сценарии, также могли бы быть достигнуты при меньших объемах инвестиций и эксплуатационных затрат ОПЕХ (соответственно на 11 и 7,5%), причем показатель NPV был бы при этом на 16% выше (28,5 вместо 24,5 условных единиц).

6.5. Заключение

Учет факторов рисков, которыми руководствуются в своей деятельности ведущие компании в области ИКТ, позволяет выявить различные типы существующих угроз, которые могут усложнить производственную деятельность как производителей, так и потребителей высокотехнологичных продуктов.

Форсированная цифровизация без должного учета рисков, связанных с киберугрозами, и без обеспечения высокого уровня кибербезопасности, приводит к серьезным экономическим потерям.

Рассмотренный игровой подход позволяет получить предварительную оценку целесообразных затрат на кибербезопасность для участников олигополистических рынков, а также разрабатывать сценарии поведения участников рынка с учетом нестационарной во времени интенсивности кибератак.

Использование введенного в главе индекса восприятия фирмами интенсивности кибератак упрощает формирование гипотез, закладываемых в основу сценариев развития рынков.

Список литературы к главе 6

1. Варшавский Л.Е. Кризис финансовой системы и эволюция товарных рынков // Прикладная эконометрика. 2010. № 1 (17). С. 30–44.
2. Варшавский Л.Е. Анализ факторов риска, связанных с информационно-коммуникационными технологиями // Концепции. 2017. № 1 (36). С. 37–47.
3. Варшавский Л.Е. Моделирование развития олигополистических рынков при наличии киберугроз // Труды Института системного анализа. 2020. Т. 70. Вып. 2. С. 21–29.
4. Качалов Р.М. Управление хозяйственным риском. М.: Наука, 2002.
5. Швецов А.Н. «Информационное общество». Теория и становление в мире и в России. М.: URSS, 2011.
6. Charette R. US Army's Future Combat Systems Program Formally Terminated. URL: <http://spectrum.ieee.org/riskfactor/computing/it/us-army-future-combat-systems-program-formally-terminated>
7. Monuments to Failure. URL: <http://spectrum.ieee.org/static/monuments-to-failure>
8. Sudhamalal H. Capitalizing on IT infrastructure services for an effective IT risk management in banks. URL: www.wipro.com
9. Варшавский Л.Е. Проблемы повышения энергоэффективности аппаратных средств в области информационных технологий // Труды Института системного анализа Российской академии наук. 2013. Т. 63. Вып. 3. С. 3–19.
10. URL: <https://www.enterprisetech.com/2017/04/09/train-locomotives-without-wheels-moats-cyber-security-innovations/>
11. URL: <http://searchcio.techtarget.com/news/450412805/The-IoT-explosion-is-upon-us-whither-CIOs>
12. URL: <https://easaily.com/news/2016/02/02/kiberataki-povzrosleli-hakery-nachali-otklyuchat-svet-v-gorodah>

13. URL: http://www.rbc.ru/technology_and_media/13/11/2016/5825cf889a79475b671ff971
14. URL: <https://www.siliconrepublic.com/machines/internet-meltdown-mirai-botnet>
15. URL: <http://www.csoonline.com/article/2989302/cyber-attacks-espionage/average-business-spends-15-million-battling-cybercrime.html>
16. URL: <https://www.gartner.com/en/newsroom/press-releases/2018-08-15-gartner-forecasts-worldwide-information-security-spending-to-exceed-124-billion-in-2019>
17. URL: <http://cybersecurityventures.com/cybersecurity-market-report/>
18. URL: <http://www.rbc.ru/newspaper/2016/11/23/5833f7289a79476da86aec0d>
19. URL: <http://www.bankinfosecurity.com/assessing-cost-cybersecurity-a-8529>
20. Villasenor J., Tehranipoor M. The Hidden Dangers of Chop-Shop Electronics. URL: <http://spectrum.ieee.org/semiconductors/processors/the-hidden-dangers-of-chopshop-electronics>
21. URL: [https://www.cnn.com › 2018/02/22 › cybercrime-pandemic-may-have-cost-\\$600billion](https://www.cnn.com › 2018/02/22 › cybercrime-pandemic-may-have-cost-$600billion)
22. URL: http://www.ng.ru/economics/2019-11-05/1_7718_hakers.html
23. Варшавский Л.Е. Использование методов теории управления для формирования рыночных структур // Компьютерные исследования и моделирование. 2014. Т. 6. № 5. С. 839–859.
24. Basar T., Olsder G.J. Dynamic Noncooperative Game Theory. London: New York: Academic Press, 1995.
25. Dockner E.J., Jorgenson S. et al. Differential Games in Economics and Management Science. Cambridge: Cambridge University Press, 2000.
26. Варшавский Л.Е. Прогнозирование динамики показателей олигополистических рынков высокотехнологичных производств с использованием методов операционного исчисления // Труды Института системного анализа. 2019. Т. 69. Вып. 2. С. 3–16.
27. Варшавский Л.Е. Модели и методы расчета динамики ввода производственных мощностей // Экономика и математические методы. 1987. Т. 23. Вып. 3. С. 456–467.

ГЛАВА 7.

СОЦИАЛЬНО-ЭКОНОМИЧЕСКИЕ ФАКТОРЫ РИСКА ИКТ ДЛЯ ОБЩЕСТВА

7.1. Введение

Развитие и использование информационно-коммуникационных технологий привело к революционным преобразованиям в экономике [1]. Наметился прогресс в таких областях ИКТ, как машинное обучение, большие данные, робототехника и др. направлениях, объединяемых в рамках понятия «искусственный интеллект», возникшего на заре развития кибернетики в 1950-х гг. Однако после финансового кризиса 2008–2009 гг. темпы распространения ИКТ превысили ранее мыслимые пределы. Сейчас, фактически, открыт «ящик Пандоры»: в отличие от других областей науки и техники (ядерной физики, авиации и др.), алгоритмы ИИ доступны для многих, что, с одной стороны, позволяет даже студентам университетов и колледжей создавать свои компании, а с другой стороны – повышает риск осуществления анти-общественных действий [2].

Причины спешки со стороны инвесторов и фирм-разработчиков новых продуктов и технологий, привыкших к быстрой прибыли, достаточно прозрачны: успеть занять желаемое место в конкурентной борьбе, сэкономить расходы на тестирование инноваций, добиться роста фондовых индексов и одобрения акционеров и др. [3].

Как отмечал один из видных специалистов в области ИКТ, внесший значительный вклад в развитие и распространение Интернета, М.Л. Дертузос, «огромные открывающиеся перед нами возможности... появляются скорее благодаря достигнутому уровню развития техники, а не из-за настоятельных нужд людей» [4, с. 12]. Таким образом, в отдельные периоды времени имеет место совпадение интересов разработчиков инноваций и инвесторов. Успеху продвижения на рынок непроверенных инноваций способствует также и неизбежное запаздывание подготовки и принятия нормативно-правовой базы, регулирующей условия распространения инноваций. Как отмечается в монографии [5], посвященной финансовой инженерии, «...адаптивные механизмы законодательства могут оказаться медленными, и в течение некоторого времени сохраняются возможности для использования его несовершенств».

На угрозы, связанные с форсированным распространением ИКТ, давно указывали основоположники и ведущие специалисты в области кибернетики. Однако сверхскоростная диффузия ИКТ происходит в условиях игнорирования предостережений основоположников кибернетики и computer science о проблемах и опасностях,

связанных с их использованием. Цель настоящей главы состоит в исследовании социально-экономических факторов риска ИКТ для общественного развития. Материал данной главы основан на статьях автора [6, 7].

7.2. Влияние технологий искусственного интеллекта на развитие общества

7.2.1. Прогресс в развитии искусственного интеллекта (ИИ)

Достижения в области элементной базы во многом стимулировали повышение интереса к исследованиям в области ИИ. Так, если ИИ, как научная область зародился в начале 1950-х гг., то научные методы и коммерческие приложения ИИ получили бурное развитие в последние 10 лет. Только в 2015 г. технологические компании (включая Google и Microsoft) потратили 8,5 млрд долл. на ИИ, что в 4 раза больше, чем в 2010 г. Вместе с тем, в [8] отмечается, что прогресс в области ИИ достигнут в 6 технологиях ИИ, которые авторы относят к «узким» (компьютерное зрение, машинное обучение, работа с естественными языками, Интернет вещей (IoT), робототехника и логические выводы (англ. reasoning)).

В работе [9] проведен анализ влияния возможных направлений ИИ на инновационную деятельность. Исследовались возможности, связанные с использованием трех направлений ИИ: робототехникой, символьными системами (symbolic systems) и глубинным обучением (deep learning). По мнению авторов, наиболее позитивное влияние на изменение характера инновационного процесса могут оказать научные исследования в области глубинного обучения.

В отмеченной работе проведен библиометрический анализ публикационной и патентной деятельности в области ИИ. Так, анализ 98 124 **публикаций** в области ИИ из базы публикаций в Web of Science за период с 1955 г. по 2015 г. показал, что значительная часть работ связана с машинным обучением (61,4%) и робототехникой (21,4%). После 2009 г. наметился также бурный рост публикаций в области применения методов глубинного обучения (в настоящее время их доля в общем числе публикаций по ИИ составляет 50%).

В то же время наибольшее число **патентов** за исследовавшийся долгосрочный период было связано с робототехникой (40%), а на остальные направления ИИ приходилось примерно равное количество патентов. В последние годы отмечается более быстрый рост числа патентов в области глубинного обучения по сравнению с другими направлениями ИИ.

Важный вывод, содержащийся в ряде исследований в области оценки влияния ИИ на экономическое развитие, относится к тенденциям локализации производства из-за его удешевления в результате автоматизации и, как следствие, к изменению характера мировой торговли (включая объемы и направления потоков товаров и услуг). Отмечается также возможность возрастания числа более мелких предприятий, приближенных к потребителю, и, как следствие этого, существенно более быстрой и эффективной адаптации производства к спросу [8, 10].

Следует однако отметить, что, дальнейший успех в области использования такого направления ИИ, как машинное обучение, связан, по мнению ряда специалистов, не с алгоритмическим обеспечением, а с логистикой, характеризующей данные, т.е. с их качеством, стандартизацией, доступностью. На ключевое значение доступности и обмена данными между специалистами в разных областях академического и коммерческого секторов для эффективной деятельности в области ИИ указывают многие специалисты. В эпоху 4-й промышленной революции недопущение к данным может служить существенным искусственным барьером и в рыночной конкуренции. В связи с этим, предлагается осуществлять государственное регулирование процессов распределения и обмена данными^{xv}.

В настоящее время начался переход к третьему этапу работ в области ИИ (Third Wave AI). Так Агентство DARPA запустило программу «Исследование искусственного интеллекта» (ИИИ, англ. Artificial Intelligence Exploration). В отличие от предыдущих этапов исследований в области ИИ, основанных на использовании большого объема данных для обучения, а также заранее заданных правил принятия решений, в ходе работ по этой программе должны быть разработаны самообучающиеся системы, способные сами генерировать решения на основе малых выборок данных [11].

Вместе с тем, технологии ИИ пока не дают 100%-й гарантии в правильности вырабатываемых решений. Так, оказывается, что методы глубинного обучения весьма чувствительны к небольшим изменениям анализируемых объектов, что может приводить не только к серьезным ошибкам, но и способствовать умышленному изменению характеристик объектов для достижения противоправных целей. Кроме того, крайне затруднена идентификация ошибок, сделанных в ходе принятия решений на основе обучения нейронных сетей [12].

^{xv} URL: <https://www.datanami.com/2018/01/04/2018-will-data/> (дата доступа 09.01.2018).

7.2.2. Глобальная конкуренция в области ИКТ

В связи с оптимистическими ожиданиями инвесторов относительно революционного переворота в экономике, благодаря достижениям в области ИКТ, ведущие страны мира начинают разрабатывать национальные программы развития ИИ и перехода к цифровой экономике. При этом США и Китай претендуют на всеобъемлющее лидерство в области ИИ. Некоторые аналитики сравнивают складывающуюся в мире конкуренцию в области ИИ с соревнованием между СССР и США в освоении космоса в годы холодной войны^{xvi}. По-видимому, первым этапом конкуренции в области ИКТ явилась гонка производительности суперкомпьютеров.

7.2.2.1. Конкуренция за производительность суперкомпьютеров

Особенностью текущего десятилетия стала утрата единоличного лидерства США в производительности суперкомпьютеров. Так, как отмечалось в главе 4, в 2010–2017 гг. большую часть времени наиболее производительными суперкомпьютерами из списка Top 500 были китайские. Несмотря на то, что в 2018 г. США удалось вернуть лидерство (производительность самого мощного суперкомпьютера Summit по показателю R_{\max} составляла 143,5 Пфлопс), в том году в списке Top 500 состояло лишь 109 американских суперкомпьютеров, в то время как 227 китайских и 31 японская [13]^{xvii}.

Следует отметить, что создание производительных суперкомпьютеров в Китае осуществляется в рамках пятилетнего плана, с целью преодоления отставания в уровне экономического развития от уровня развитых стран, а также с целью обеспечения собственных разработок в области высоких технологий. Последнее вызывает озабоченность в США, так как в стремлении к обладанию Китаем собственными высокими технологиями там усматривается угроза повышения военного потенциала этой наиболее населенной страны мира [14]. Вместе с тем, нельзя исключать того, что в ближайшем будущем лидерство в области наиболее производительного суперкомпьютера сохранится у американцев, которые уже успешно реализуют ряд амбициозных проектов, связанных с созданием эксафлопсного суперкомпьютера.

7.2.2.2. Национальные планы и программы развития ИИ

Один из первых развернутых документов, в котором предлагалась стратегия развития ИИ с учетом его позитивных и негативных последствий явился доклад Национального Совета по науке и технологиям США «Подготовка к будущему ИИ»

^{xvi} URL: <https://www.csis.org/analysis/machine-intelligence-we-need-national-strategy/> (дата доступа 30.07.2018).

^{xvii} URL: <https://www.top500.org/news/lists/2018/11/press-release/> (дата доступа 30.11.2018).

(«Preparing for the Future of Artificial Intelligence»), подготовленный в 2016 г. Большое место в докладе уделяется вопросам влияния ИИ на экономику, в частности, на повышения производительности, а также на рынок труда. В докладе отмечается, что в течение ближайших 20 лет не следует ожидать, что ИИ будет превосходить человеческие возможности для решения все большего числа задач. Вместе с тем, указывается, что ИИ и автоматизация будут способствовать процветанию американской экономики, создавая, однако, трудности для некоторых категорий трудящихся (операторов, клерков, турагентов, рабочих на сборочных линиях). Приводятся данные о том, что в течение ближайшего десятилетия от 9 до 47% рабочих мест может находиться под риском исчезновения, что, в свою очередь, усилит экономическое неравенство. В связи с этим, Совет предложил разработку и реализацию государственных мер по стимулированию перехода к ИИ и смягчению неблагоприятных последствий от развития ИИ.

Первая мера (Invest in and develop AI for its many benefits) предполагает увеличение государственного инвестирования в исследования и разработки в области ИИ. Однако при этом потребуются увеличить численность специалистов по техническим наукам и математике (STEM). Вторая мера (Educate and train Americans for jobs of the future) состоит в удовлетворении потребности рынков в работниках с требуемыми навыками. Для этого предлагается увеличить финансирование образовательных учреждений, а также обеспечить помощь работникам в переквалификации в нужном направлении. Наконец, третья мера (Aid workers in the transition and empower workers to ensure broadly shared growth) направлена на создании условий для обеспечения работой и повышения зарплат квалифицированных работников, совершенствование системы социальной защиты и системы здравоохранения. [15].

Необходимость более широкого вовлечения государства в развитие ИИ активно лоббируется и ведущими американскими компаниями в области ИКТ, которые начинают понимать не только позитивные, но и негативные последствия использования ИИ. Показательными являются обращения представителей компании Intel к правительству относительно поддержки проектов Закона об открытых правительственных данных (OPEN Government Data Act), а также Закона о влиянии ИИ на рабочие места (AI JOBS Act). Первый из упомянутых проектов должен способствовать допуску специалистов по машинному обучению к государственным базам данных, являющихся общенациональным ресурсом и формируемым за счет средств налогоплательщиков. Цель второго предложения – смягчить неблагоприятные последствия возможной потери работы трудящимися в результате использования ИИ. Представляет интерес предложение Intel обеспечить государственную поддержку прозрачности и объяснимости (explainability) алгоритмов программного

обеспечения ИИ, которое для многих потребителей, включая крупный бизнес, являются «черным ящиком». Такая мера, по мнению специалистов Intel, поможет устранить информационное неравенство, дискриминацию и вред пользователям средств ИИ [6].

После появления американского доклада о будущем ИИ в США, в середине 2017 г. Государственным Советом КНР был опубликован план по развитию ИИ следующего поколения (Next Generation Artificial Intelligence Development Plan) на период до 2030 г. В плане установлены 3 этапа развития ИИ в Китае. На первом этапе, к 2020 г. технологии ИИ должны достичь мирового уровня и стать драйвером экономического развития страны. На втором этапе, к 2020 г. должны были быть получены прорывные результаты в теории ИИ. ИИ должен стать драйвером промышленного обновления и экономической реструктуризации. Должны быть заложены основы для построения интеллектуального общества в стране. К 2030 г теория, технологии и применения ИИ должны достичь глобального уровня, Китай должен стать глобальным центром ИИ. Должен быть достигнут прогресс в построении интеллектуальной экономики и общества.

К числу важных задач в плане относится создание надежной и безопасной инфраструктуры ИИ, обеспечение интеграции гражданских и военных исследований в области ИИ (Enhance AI civil-military integration). С этой целью предполагается наладить совместную работу крупных ИТ компаний с национальными лабораториями. Комплекс мер, которые должны способствовать выполнению плана, включает: 1) формирование законов, регулирующих актов и этических норм для развития ИИ; 2) улучшение общей политики стимулирования развития ИИ; 3) установление стандартов для технологий ИИ, а также системы интеллектуальной собственности; 4) регулирование безопасности ИИ и установление системы оценки ИИ; 5) обеспечение подготовки кадров для ИИ; 6. активизация деятельности по популяризации ИИ [16].

Стратегия развития ИИ разработана и в Канаде (Pan-Canadian Artificial Intelligence Strategy). Основными целями этой стратегии являются: увеличение численности высококвалифицированных исследователей и аспирантов в университетах в области ИИ; установление взаимосвязи между ведущими канадскими центрами в области ИИ в Эдмонтоне, Монреале и в районе Торонто-Уотерлоо; обеспечение глобального интеллектуального лидерства в области оценки экономических, правовых и этических последствий, связанных с достижениями ИИ; поддержка национальных исследований в области ИИ.

Интересно отметить, что ведущими канадскими отраслями в области использования ИИ считаются маркетинг, финансовые услуги, здравоохранение с фармацевтикой, кадровые, а также социальные службы [6].

Другие страны также не отстают в желании занять достойное место в ходе происходящих изменений (перехода к 4-й промышленной революции). Так, в ЕС разработаны меры по интенсификации инвестиций в развитие ИИ, подготовке социально-экономических изменений и обеспечению этических и правовых норм в указанной области [6]^{xviii}.

7.3. Влияние ИИ на рынок труда

7.3.1. Оценки академических и консалтинговых организаций

К исследованию проблемы оценки влияния ИИ, автоматизации и, особенно роботизации, на уровень занятости в экономике, активно обсуждаемой в последние годы на Давосских экономических форумах под лозунгом перехода к 4-й промышленной революции, подключился ряд солидных академических и консалтинговых организаций.

Так, в работах D. Acemoglu с сотрудниками (см., например, [17]) отмечается, что за последние 2–3 десятилетия, несмотря на технологические достижения, в экономике развитых стран происходило снижение производительности труда и заработной платы в реальном исчислении. В связи с этим, автор выделяет 2 вида технологий: перспективные (англ. *enabling*) и заменяющие (англ. *replacing*), которые по-разному воздействуют на рынок труда. Первый вид технологий способствует повышению производительности труда, заработной платы и спросу на рабочую силу (за счет расширения возможностей естественного интеллекта). Второй вид технологий приводит к снижению спроса на рабочую силу и ее высвобождению за счет автоматизации и роботизации производств (в первую очередь, низкоквалифицированной рабочей силы). Опираясь на исторические аналогии (в частности, связанные с последствиями внедрения в XIX в. фабричных форм производства, появления паровых двигателей и др.), авторы отмечают, что перспективные технологии могут приводить к эффективности производства и росту спроса на рабочую силу не сразу, а в течение достаточно длительного периода.

Интенсивные исследования в рассматриваемой области проводятся и в таких крупных консалтинговых компаниях, как McKinsey, PwC и др. Так, в соответствии с одним из прогнозных сценариев исследователей McKinsey, в 2016–2030 гг. за счет автоматизации производства будет вытеснено 15% занятых (400 млн человек). Однако еще большему числу работников, как считают в McKinsey, повезет. Так, будет

^{xviii} URL: <http://ai-europe.eu/the-european-commission-outlines-an-approach-to-boost-investment-and-set-ethical-guidelines/> (дата доступа 04.05.2018).

создано от 555 млн до 890 млн новых рабочих мест, что в глобальном разрезе составит 21–23% от всей рабочей силы. Как и следует ожидать, в соответствии с прогнозами в первую очередь пострадают работники, связанные с физическим трудом, а также с малоквалифицированной переработкой данных (data processing). Сокращение не затронет представителей таких профессий, которые, по мнению исследователей McKinsey трудно автоматизировать: менеджеров, учителей, нянь, а также садовников, водопроводчиков и других профессионалов, которые работают в трудно предсказуемых условиях (who work in unpredictable physical environments).

Вместе с тем, автоматизация будет оказывать давление на уровень заработной платы. Это связано с тем, что ряд высокооплачиваемых специальностей (например, в машиностроении, в бухгалтерском деле) исчезнет. И хотя заработная плата высококвалифицированных медицинских работников и других профессионалов будет расти, уровень заработной платы большей части работников (учителей, нянь и др.) будет невысоким.

Доля представителей физического и ручного (manual) труда в США и Западной Европе в общей численности занятых уменьшится, по оценкам исследователей McKinsey, с 31% в 2016 до 25% в 2030 г. Вместе с тем, в течение ближайших 3 лет сохранится высокий спрос на программистов и высококвалифицированных специалистов в области информационных технологий (ИТ) [18]^{xix}. Эти оценки согласуются с мнением других исследователей, которые считают, что труд ученых, технологов, инженеров и представителей математических специальностей (science, technology, engineering and math jobs – STEM) в ближайшее время не удастся автоматизировать [19].

Результаты отмеченного исследования McKinsey интересно сопоставить с выводом К. Маркса, сделанным в XIX в.: «...чрезвычайно возросшая производительная сила в отраслях крупной промышленности... дает возможность... воспроизводить все большими массами старинных домашних рабов под названием «класса прислуги», как, например, слуг, горничных, лакеев и т.д. ... Что за превосходный результат капиталистической эксплуатации машин!» [20, с. 456, 457].

Близкие выводы к прогнозам других консалтинговых организаций содержатся в более позднем докладе Центра новой экономики и общества (Centre for the New Economy and Society) Всемирного Экономического Форума (ВЭФ), подготовленном на основе данных опроса совокупности компаний, в которых в целом занято более 15 млн сотрудников. В докладе приведены оценки численности и пропорций работников тех специальностей, которые признаны избыточными (redundant) и перспективными (emerging). К первой группе отнесены, в частности, разного рода

^{xix} Неясно, однако, как спрос на этих специалистов изменится в более далекой перспективе.

клерки, бухгалтеры, финансовые аналитики механики, машинисты и даже юристы. Во вторую группу включены в основном специалисты в области ИКТ и ИИ (специалисты в области анализа больших данных, машинного обучения, робототехники, менеджеры, и др.).

Данные опросов, проведенных до распространения COVID-19, показывают, что больше всего компаний (85%) собиралось использовать в 2022 г. технологии больших данных и бизнес-аналитику. Существенно меньшее число компаний собиралось использовать роботов разных видов (от 19 до 38% в зависимости от видов и отрасли деятельности).

Около 50% компаний считает, что в результате автоматизации численность их работников сократится. Доля работников избыточных специальностей уменьшится с 31% в 2018 до 21% в 2022 г., а перспективных за тот же период – увеличится с 16 до 27%. 54% работников придется переобучаться, причем делается оптимистичное предположение о том, что переобучение (reskilling) займет не более 1 года [21].

7.3.2. Моделирование динамики занятых и безработицы в результате вытеснения труда ИИ

Для выяснения вопроса о том, к каким последствиям на занятость населения на условной территории (в стране) может привести вытеснение труда средствами ИИ, нами разработана следующая агрегированная модель:

$$x_t = (1 - \lambda)x_{t-1} + y_t + u_t; \quad (7.1)$$

$$y_t = \gamma\lambda(1 - \mu)x_{t-1} + \mu y_{t-1}; \quad (7.2)$$

$$U_t = \sum_{\tau=0}^t (\lambda x_\tau - y_\tau), \quad (7.3)$$

где x_t – общая численность занятых на территории; y_t – численность переобученных работников, принятых на работу (после переквалификации); u_t – величина притока новых кадров (после образовательных учреждений, в расчетах принято, что $u_t = 0,01 x_0$, $t = 0, 1, 2, \dots$); U_t – численность безработных (принято, что $U_0 = 0$, $y_0 = 0$). Коэффициенты λ, γ, μ характеризуют соответственно долю сокращаемой рабочей силы в общей ее численности, долю переквалифицирующейся сокращенной рабочей силы, а также инерционность в устройстве на новую работу (средний период устройства на новую работу на основе соотношения (7.2) определяется как

$$T = 1 + \frac{\mu}{1 - \mu} \text{ лет}).$$

Расчеты показывают, что в случае умеренных темпов вытеснения рабочей силы средствами ИИ ($\lambda = 0,03$), даже при высокой доле переквалифицирующихся

кадров ($\gamma = 0,75$), уровень безработицы $UN_t = U_t/(x_t + U_t)$ будет расти как при быстром трудоустройстве уволенных ($\mu = 0,35$, $T = 1,5$ года), так и при медленном трудоустройстве (года $\mu = 0,7$, $T = 3,3$). Через 5 лет после начала внедрения ИИ UN_5 составит в первом случае 5%, а во втором – 8%. Через 10 лет уровни безработицы для рассматриваемых вариантов составят соответственно 8 и 11%. Безработицы не удастся избежать и при снижении интенсивности замены занятых λ (с 0,03 до 0,01) (ее уровень составит через 5 лет 3,5%, а через 10 лет возрастет до 5%).

7.4. Социальные факторы риска для общества

7.4.1. Предвидения видных ученых на заре развития ИКТ

Первые предостережения со стороны основоположников кибернетики о рисках, связанных с тотальной автоматизацией и использованием искусственного интеллекта были сделаны еще в 1950–1960-х гг. Так, основоположник кибернетики Н. Винер (к сожалению, имя этого выдающегося ученого неизвестно многим современным студентам, изучающим и изучавшим информатику) отмечал, что кибернетические устройства могут быть нашими помощниками, «но при условии, что наша честь и разум будут удовлетворять требованиям самой высокой морали». Расплата за ошибки в создании и использовании таких устройств и систем «...еще больше возрастет, когда автоматизация достигнет полного размаха» [22].

Уже 50 лет назад выдающихся ученых волновали возможные последствия непродуманного использования вычислительной техники, связанные с ограничением прав и свобод граждан, и, прежде всего, права на неприкосновенность личной жизни, с риском возникновения информационной тирании и неравенства, деградации интеллектуального потенциала людей и всего общества. Так, известный специалист в области управления С. Бир, консультант правительства С. Альенде в Чили в 1970-е гг., задавался вопросом: разумно ли, что «большая часть достижений в области улучшения благосостояния оплачена нарушением личных свобод». И далее: «вполне возможно, что вычислительная техника будет бурно развиваться, дело будет идти к полному разрушению конфиденциальности и устранению свободы выбора, а мы даже не будем предполагать, что этот процесс происходит» [23, с. 18, 25].

В начале 1970-х гг. Президент Массачусетского технологического института Дж.Б. Визнер также предупреждал, что «...свобода индивидуума... может быть принесена в жертву ради «всеобщего блага» технического развития», причем «...особая опасность применения средств передачи информации кроется в их легкой доступности как для правительства, так и для частных лиц. Это создает угрозу нарушения

хрупкого равновесия «человек – государство, установленного конституцией и устраняющей анархию, с одной стороны, и тиранию – с другой...». И далее: «Наша задача заключается в установлении равновесия между возможностью устранения угроз обществу со стороны индивидуума и способностью общества лишать своих членов свободы и счастья» [24. с. 42, 44–45].

В докладе Комитета «ЭВМ и общество» Ассоциации вычислительных машин (председатель Комитета Д.Д. Маккракен) в 1974 г. также содержится предупреждение, что «...скоро для практической экономики и политики будут «существовать» лишь те люди, чью личность можно будет идентифицировать при помощи ЭВМ. Все остальные не будут иметь никаких прав...» [25, с. 48].

В этом докладе перечислено 13 проблем, требующих серьезного изучения, среди которых: информационные услуги для домашнего пользования; ЭВМ и деньги; ЭВМ и выборы; ЭВМ и процесс обучения; ЭВМ как метафора человеческого самопознания; ЭВМ и конфиденциальность; ЭВМ и занятость; грамотный подход к ЭВМ; подготовка специалистов в области вычислительной математики, кибернетики и обработки данных; ответственность; монополия; вычислительные машины и патенты, авторское право и коммерческая тайна; вычислительные машины и электронная передача данных.

Авторы доклада спрашивают: «Не приведет ли... легкость голосования (*электронного*) к получению огромного количества случайных мнений?.. Насколько велика опасность использования такой системы для слежки каким-нибудь диктатором?.. Насколько хорошо эта система может быть защищена от саботажа?..»; «Не приведет ли использование карманных вычислителей к возникновению нации людей, не умеющих умножать и складывать?..» [25, с.49]. Отмечается, что «...хотя применение ЭВМ создает сотни тысяч рабочих мест, оно в то же время упраздняет или изменяет множество существующих... некоторым уволенным людям будет трудно или невозможно переквалифицироваться для другой работы» [25, с. 58].

Следует подчеркнуть, что на поднятые специалистами 50–60 лет назад вопросы еще не дано адекватного ответа. Об этом свидетельствует, например, информация, содержащаяся в выпусках Risk Digest. Forum on Risks to the Public in Computers and Related Systems. ACM Committee on Computers and Public Policy, начиная с 1985 г. и до настоящего времени.

Не решены даже такие вопросы, как «кто должен отвечать за сбой машинной программы?», поставленные в 1970-е гг. Так, как отмечается в учебнике по информатике Дж. Брукшира, изданному в США в 2003 г. «...если компьютеризованная база данных выдаст ошибочную информацию о вашей низкой кредитоспособности,

неверные данные о преступлении или о превышении чекового счета, кому поверят – утверждению компьютера или вашей клятве в невиновности?...» [26, с. 60].

В последнее время все большее число ученых в области общественных наук начало исследовать подобные проблемы. Так, Kenneth Cukier and Viktor Mayer-Schoenberger, авторы известной книги о больших данных, изданной в 2013 г., предупреждают, что задача этого направления ИКТ и ИИ – информировать, а не объяснять; что использование больших данных может приводить к ошибкам, что нужно учитывать не только их мощь, но и ограничения [27, р. 40].

7.4.2. Социальные последствия форсированного развития ИКТ

Развитие ИКТ и, в частности, ИИ происходит стремительными темпами, превосходящими темпы развития базовых технологий. Фактически, открыт «ящик Пандоры»: в отличие от других областей науки и техники (ядерной физики, авиации и др.), алгоритмы ИИ доступны для многих, что позволяет даже студентам университетов и колледжей создавать свои компании^{xx}.

В связи с этим, важное значение приобретает выработка и соблюдение этических норм и правил, которыми должны руководствоваться разработчики, заказчики и пользователи ИКТ-приборов и услуг. Исследователи с тревогой предупреждают, что «неконтролируемое использование ИКТ способствует появлению новых видов преступности, среди которых наиболее опасными для общества являются терроризм, преступления против личности, в области банковской деятельности, в области охраны интеллектуальной собственности и др.». Они указывают также на опасность возникновения информационного неравенства. Как отмечают В.В. Иванов и Г.Г. Малинецкий, «...бесконтрольная централизация информации и ее дифференциация может привести к такой ситуации, когда информационный оператор сможет оказывать прямое несанкционированное воздействие на определенные слои населения» [28].

Следует отметить и такие отрицательные тенденции в использовании ИКТ, как формирование у молодых людей и детей «клипового» и «кликерного» сознания, а также фактический отказ взрослых людей от воспитательных функций в семье взамен сосредоточения на виртуальной реальности. Особенно резкий переход в этом направлении произошел в текущем десятилетии с появлением смартфонов.

Происходит обывание населения все возрастающим массивом компьютерных игр (доходы игровой индустрии в мире в 2018 г. составили по

^{xx} URL: <https://spectrum.ieee.org/tech-talk/robotics/artificial-intelligence/former-head-of-google-china-foresees-an-ai-crisis/> (дата доступа 05.09.2018).

оценкам [14] 137,9 млрд долл.), перевод его в виртуальную и дополненную реальность. По мнению ряда политиков и представителей игровой индустрии, перечисленные меры способны смягчить грядущее высвобождение рабочей силы в результате форсированного использования робототехнических устройств и ИИ, упорно отрицаемое специалистами известных консалтинговых компаний и международных организаций [5]. Однако при этом не учитываются такие риски, как возможность получения психологических травм игроков, бесконтрольное усиление их агрессивности, ведущее в итоге к совершению преступных актов и др.).

Наметилась тенденция к созданию в обществе кастовой системы (так, в КНР в ходе реализации «Программы создания системы социального кредита (2014–2020)», с использованием инструментария ИИ создается кастовая система, в соответствии с которой каждому гражданину страны, в зависимости от его поведения и «добрых дел», с помощью технологии big data будет присваиваться свой «рейтинг». Все население в зависимости от величины рейтингов будет поделено на 5 групп (начиная от группы образцовых граждан и кончая группой изгоев, которым запрещено пользоваться современными видами транспорта); система уже успешно внедрена в ряде провинций Китая и имеет своих поклонников в разных странах мира [29].

ИКТ способствовали появлению социальных сетей, благодаря которым осуществляется обмен информацией сотен миллионов людей. Однако обратная сторона этого явления состоит в том, что уже достигли достаточного уровня развития такие направления использования социальных сетей, целью которых являются тотальная слежка, провокационная деятельность для создания и поддержания в обществе напряженной обстановки, свержение неугодных режимов, вовлечение молодежи в экстремистские организации и др. [30]. Распространенным явлением стало использование персональных данных пользователей сетей в противозаконных целях. При этом необходимо отметить, что пользователи не обладают средствами контроля удаления из соцсетей своих данных: «...все держится на обещании» соцсетей [31]. Расширяется использование социальных сетей в провокационных и противоправных целях, в частности для ведения информационных войн.

Появились серьезные медицинские риски, увеличивающиеся по мере все большего охвата населения земного шара мобильной связью; недоучет медицинских рисков для общества, в частности последствий влияния значительного увеличения плотности и величины сверхвысокочастотного электромагнитного излучения на здоровье населения и на инфраструктуру.

Указанные медицинские риски особенно обостряются в последнее время, когда во многих странах мира значительные усилия компаний-провайдеров при активной поддержке правительств направлены на развертывание технологии мобильной

связи 5G. Необходимость перехода к 5G связывается с увеличением трафика, предполагаемым бурным распространением Интернета вещей IoT, и расширением масштабов обслуживания движущихся объектов и др. Для передачи сигналов с повышенной спектральной эффективностью (spectrum efficiency) западные ИКТ-компании намеряют использовать миллиметровые волны (частотой от 30 до 300 ГГц, против 800–2600 МГц для систем 1G–4G), которые, однако, поглощаются каплями дождя и частицами в воздухе, и не могут проходить через преграды. Поэтому предполагается примерно через каждые 250–500 м сооружать небольшие базовые станции (micro basestations), что резко увеличит плотность их установки (в десятки и сотни раз).

Возможные направления использования высокочастотных диапазонов (24,25–27,5 ГГц; 37–43,5 ГГц) в технологии 5G связываются с подключением пользователей с высокими пиковыми скоростями для 3D-видео, для работы и игр в облаке, для дополненной реальности, голографических звонков и др. коммерческих мероприятий, спрос на которые не очевиден [6]. Однако недостаточная маркетинговая проработка эффективности возможных направлений мобильной связи может привести к серьезным потерям компаний-провайдеров, привыкшим к быстрой отдаче и высокой марже.

Еще более важно, что не исследовано влияние значительного увеличения плотности и величины сверхвысокочастотного электромагнитного излучения миллиметрового диапазона на здоровье населения и на инфраструктуру! Вместе с тем, уже доказано, что даже электромагнитное излучение от мобильных телефонов (с используемой частотой от 800 МГц до почти 3 ГГц) влияет на электрическую активность мозга во время сна, а также на другие биологические процессы, и может способствовать возникновению онкологических заболеваний. Установлено также, что при длительном использовании мобильных телефонов возникают внутричерепные опухоли, воздействующие на нервные клетки, соединяющие ухо со структурами головного мозга. Многолетние исследования Института Ramazzini показали существенное влияние постоянно действующего высокочастотного радиоизлучения на сердечную деятельность, приводящего к невриноме сердца (schwannomas of the heart) [32, 33].

Еще большие опасения вызывает у многих биологов и биофизиков развертывание технологии 5G, которое некоторые из них считают «массивным экспериментом над здоровьем всех живых существ». В связи с тем, что миллиметровые волны поглощаются кожей, которая содержит капилляры и нервные окончания, возникает риск воздействия на нервную систему через кожные молекулярные механизмы. Ученые обращают внимание и на то, что миллиметровые волны (частотой 95

ГГц) используются в нелетальном оружии (излучение большой мощности из такого оружия может вызывать сильные ожоги) [32, 33].

Необходимо отметить, что в начале 1990-х гг., на заре распространения системы GSM (разновидность технологии 2G), для производителей оборудования мобильной связи государственными органами европейских стран были установлены лимиты по мощности излучения со стороны базовых станций и мобильных устройств. В частности, для контроля уровня излучения мобильных телефонов был введен показатель удельной величины поглощения излучения организмом человека SAR (Specific Absorbtion Rate), измеряемый в ваттах на килограмм (Вт/кг). С тех пор никаких новых требований по созданию и соблюдению безопасных для здоровья стандартов для мобильных систем более поздних поколений не было введено [34].

7.5. Социальные проблемы, связанные с распространением ИКТ в России

В нашей стране, как и в других странах мира, развивающийся рынок персональных данных является примером плохо контролируемого использования информации. По некоторым оценкам в 2017 г. почти 80% всех сайтов мирового интернета были оборудованы трекерами, которые собирают информацию о действиях пользователей мировой паутины, в основном, в рекламных целях. Через эти устройства третьи лица могут получить данные о поведении любого потребителя, зашедшего через определенный браузер (cookies). Таким образом, у третьих лиц аккумулируется информация и о персональных данных российских потребителей (о поле, возрасте, семейном статусе, профессиональных интересах, потребительских привычках и желаниях) [35].

Слежка за потребителями, пользующихся мобильными устройствами, ведется в 260 торговых центрах (ТЦ) страны. С помощью сенсоров определяется физический адрес телефонов, на основе которого идентифицируется владелец телефона. Несмотря на то, что в соответствии с требованиями Роскомнадзора, «...персональные данные должны использоваться для тех целей, для которых собраны», у потребителей нет уверенности в том, что они могут использоваться третьими лицами не по назначению [35].

Отдельного рассмотрения заслуживает проблема оценки влияния автоматизации и, особенно роботизации, на уровень занятости в экономике. По оценкам РАН-ХиГС, проведенным на основе предложенных компанией McKinsey коэффициентов потенциальной автоматизации, в России, в случае «одномоментной» роботизации, из активной хозяйственной деятельности могут быть исключены 42,13 млн человек, что

составит 49% трудоспособного населения. В первую очередь, по мнению авторов, это затронет такие профессии, как водителей, продавцов (высвобождение может коснуться 98% занятых), грузчиков (72%), бухгалтеров и экономистов (43–94%) [36].

Пока ситуацию смягчает то, что в нашей стране установлено существенно меньше промышленных роботов, чем в развитых странах (4000 против 130 тыс. в Германии). Однако для повышения производительности труда (в России она вдвое ниже, чем в среднем по странам ОЭСР) и конкурентоспособности отечественной промышленности потребуются существенно увеличить уровень роботизации. Некоторые исследователи даже отмечают, что «...эффективность и производительность труда в России приносится в жертву социальной стабильности» [37]. Вместе с тем, специалисты отмечают, что наиболее эффективно осуществляется роботизация в тех крупных компаниях, в которых возможности и производительность роботов сочетаются с аналитическими способностями человека [6].

Для смягчения негативных последствий от достаточно стихийного распространения ИКТ необходимо квалифицированное правовое регулирование со стороны государства. С этой целью Медиа-коммуникационный союзом (МКС), в состав которого входят крупнейшие в стране телекоммуникационные компании и медийные холдинги, с конца 2017 г. начата разработка Инфокоммуникационного кодекса (с участием Boston Consulting Group (BCG)). Этот всеобъемлющий документ по мысли авторов должен заменить ныне действующие в стране правовые акты. Необходим тщательный анализ последствий решений и предложений, рекомендуемых специалистами по ИКТ и, в частности, необходима разработка адекватных моделей и методов тестирования рисков, обусловленных использованием новых ИКТ [6].

Представляется целесообразным обеспечить гармонизацию нормативно-правовых актов, относящихся к регулированию ИКТ. Крайне важно в самое ближайшее время разработать нормативно-законодательную базу и кодекс этических норм и правил, которыми должны руководствоваться разработчики, заказчики и пользователи ИКТ-приборов и услуг [6].

7.6. Проблемы и риски, связанные с созданием «умных» городов

Одна из характерных особенностей территориального размещения населения нашей планеты в последнее столетие связана с бурным развитием процессов урбанизации. Эти процессы сопровождаются стремительным ростом мегаполисов (метрополисов). Число мегаполисов с населением свыше 10 млн чел. увеличилось с 1 в 1950 г. до 19 в 2000 г. и до 34 в 2020 г. Всего по данным ООН, в мире в настоящее время насчитывается 1934 мегаполиса с населением свыше 300 тыс. чел. [38].

В настоящее время в мегаполисах проживает 2592 млн чел., что более, чем в 8 раз превышает уровень 1950 г. По оценкам ООН в 2035 г. этот показатель может составить 3472 млн чел. При этом в мегаполисах будет сосредоточено соответственно 39% общего и 62,5% городского населения всего мира, а также значительная часть производительных сил планеты. Так, по данным Брукингского института, в 2016 г. на 300 наиболее крупных мегаполисов приходилось свыше 24% занятых и 49% мирового ВВП. Эта группа мегаполисов характеризуется также значительно более высокими темпами роста своей экономики по сравнению с мировой (примерно на 30%) [39].

Перемещение огромных масс людей в мегаполисы и города, сопровождающееся резким повышением плотности населения, требует решения большого числа проблем, связанных с обеспечением их жизнедеятельности. В настоящее время ключевая роль в решении проблем мегаполисов отводится ИКТ.

7.6.1. Определения «умного» города

Ввиду бурного развития урбанизации, во многом связанного с углублением процессов глобализации, а также с внедрением информационно-коммуникационных технологий (ИКТ), в конце XX в. была выдвинута концепция «умного» города.

Вначале она в значительной мере опиралась на концепцию Третьей промышленной революции, предложенную американским экономистом и экологом Дж. Рифкиным и поддержанной европейским Союзом (ЕС) и ООН. Суть ее состояла в том, что, необходим переход от традиционных централизованных методов управления экономикой к горизонтальным взаимодействиям между бизнес-структурами и, непосредственно, между потребителями, которые, по мере развития «зеленых» технологий и ИКТ смогли бы принять участие в децентрализованном производстве и распределении разного рода городских ресурсов [40]. В дальнейшем под «умным» городом стали понимать объект с более широкими свойствами. Появилось много формулировок, которые охватывали достаточно широкий спектр близких проблем, связанных с устойчивым развитием городов.

В частности, одно из удачных определений дано сотрудниками компании Deloitte. Так, в соответствии с этим определением, город следует считать «умным», когда инвестиции в человеческий и социальный капитал, традиционную инфраструктуру, прорывные технологии обеспечивают экономический рост, высокий уровень жизни, а также эффективное использование природных ресурсов путем participatory governance).

Под прорывными технологиями в Deloitte понимаются такие технологии ИКТ, как большие данные, Интернет вещей (IoT), облачные вычисления, мобильная связь,

робототехника, искусственный интеллект, разного рода цифровые платформы для социальных взаимодействий, краудсорсинг и др. [41].

Другое определение «умного» города, ориентированное, главным образом, на продвижение расширившихся возможностей ИКТ, дано Международным союзом связи (ITU). Так, под «умным» устойчивым городом понимается инновационный город, в котором используются ИКТ и другие средства, направленные на улучшение качества жизни, повышение эффективности городских операций и сервисов, конкурентоспособности, и в котором обеспечиваются потребности текущего и будущих поколений граждан в экономическом, социальном и экологическом аспектах [42].

Для оценки прогресса в реализации проектов создания «умных» городов используются рейтинги различных организаций. Так, в одной из крупнейших телекоммуникационных компаний Ericsson рассчитывается сводный индекс **Networked Society City Index** на основе показателей деятельности 15 социально-экономических сфер и показателей, выделенных в 2 группы: 1) здравоохранение, образование, социальное участие (Social inclusion), производительность, конкурентоспособность, уровень загрязнения, изменение климата, усилия по смягчению экологической ситуации; 2) качество связи, ее доступность, тарифы на связь, цены передачу данных, использование ИКТ как на индивидуальном уровне, так и в городском хозяйстве и в рыночной среде [43].

Одной из первых организаций, осуществляющих конкурс наиболее подготовленных городов к статусу «умного города» является Intelligent Community Forum (**ICF**). Ежегодно **ICF** выделяет 7 городов, имеющих наибольшие успехи в реализации ИКТ-проектов по данной тематике.

Достаточно авторитетным считается также рейтинг «умных» городов мира **Smart Cities Index**, разрабатываемый Easypark, в котором также основное внимание уделяется использованию ИКТ в городской жизни (наличие 4G и 5G, общественного Wi-Fi, интенсивность использования смартфонов, онлайн-доступ к государственным услугам, интеллектуальных систем управления транспортом и др.). Индекс охватывает показатели, распределенные в 19 групп: 1) уровень развития каршеринг сервисов; 2) уровень интеллектуальности парковочных пространств; 3) уровень загруженности дорог; 4) степень удовлетворенности населения работой общественного транспорта; 5) показатель развития умных зданий (ВВП на единицу энергопотребления и др.); 6) степень экологичности энергетики; 7) качество утилизации отходов; 8) оценка уровня загрязнения окружающей среды; 9) уровень политической активности граждан; 10) степень цифровизации системы государственного управления (трафик интернет-сайтов местной администрации и др.); 11) оценка планирования городской среды; 12) уровень образования (количество университетов города,

входящих в состав 10 лучших университетов и др.); 13) оценка бизнес-экосистемы; 14) оценка качества 4G LTE; 15) оценка скорости Интернета; 16) количество точек доступа к бесплатному Wi-Fi ; 17) показатель охвата населения города смартфонами; 18) оценка уровня жизни в городе (средняя сумма, потраченная на общественное питание, одежду, аренду, транспорт, средняя чистая заработная плата с учетом уровня ВВП на душу населения и др.); 19) экспертная оценка [44].

Кроме того, рейтинги «умных» городов выпускают институт McKinsey, компания PwC (индекс готовности городов), а также Центр глобализации и стратегии IESE (ежегодный рейтинг **Cities in Motion Index (CIMI)**). Рейтинг **CIMI** считается одним из наиболее представительных – в 2018 г. охватывал 180 городов из 80 стран мира [45]. Другим представительным индексом является **IMD-SUTD Smart City Index (IMD)**, подготавливаемым IMD World Competitiveness Center, Lee Kuan Yew Centre for Innovative Cities, а также IMD Smart City Observatory.

Следует, однако, отметить несопоставимость рейтингов различных организаций, что связано, в частности с разной методикой обработки показателей, характеризующих социально-экономическую ситуацию в городах. В разных системах подсчета рейтингов первые места занимают разные города. Например, в соответствии с индексом **CIMI**, Нью-Йорк, Лондон и Сан-Франциско^{xxi} входят в группу из 10 лидеров. В то же время по индексу **IMD**, сформированному в текущем году, эти города занимали соответственно 10-е, 15-е и 27-е место [46, 47].

Такая несопоставимость индексов помогает администрациям городов поддерживать в ряде случаев городской имидж с надеждой на то, что даже некоторые повышенные рейтинги, или их благоприятная динамика, привлекут могущественные компании в области ИКТ и крупных инвесторов к созданию или обновлению городской инфраструктуры и среды. В этом отношении показательно стремление администраций европейских городов, на которые, одновременно с урезанием городских бюджетов после кризиса 2008–2009 гг., была перенесена большая ответственность за социально-экономическое состояние и развитие городов, получить гранты в рамках Рамочных программ Европейской Комиссии [48]. Вместе с тем, со стороны общественности и ряда исследователей появилась критика того, что в подготовленных и реализуемых проектах «умных» городов игнорируются потребности граждан, что в этих проектах не предусмотрено участие граждан в управлении городами. Провозглашаемый рядом идеологов «умных» городов принцип активного вовлечения граждан в управление (самоуправления) фактически, в угоду крупным компаниям и

^{xxi} Интересно отметить, что в Сан-Франциско уже в 2017 г. свыше 40% потреблявшейся в городе энергии вырабатывалось за счет возобновляемых источников, что позволило уменьшить на 25% выбросы парниковых газов)

администрации, заменен традиционным принципом «Top-down governance» (управление сверху вниз).

Некоторые исследователи отмечают 5 признаков, присущих разрабатываемым проектам «умных» городов: 1) поддержка неолиберального направления в экономике, основанного на приоритете показателей экономического роста в городах; 2) фокусировка и опора на наиболее зажиточной части населения; 3) ущемление прав и маргинализация граждан; 4) игнорирование проблем охраны окружающей среды; 5) нежелание изменить сложившиеся стереотипы потребительской культуры [49].

Одним их характерных примеров является попытка продвижения концепции «умного» города в Филадельфии через проект **Digital On-Ramps Initiative**, разработанный при участии компании IBM. Так, предполагалось, за счет использования мобильной связи и смартфонов, подготовить 500 тыс. жителей города к работе в условиях новой, цифровой экономики и экономики знаний. Реализация проекта подверглась серьезной критике, так как в нем игнорировались накопившиеся за многие годы социальные проблемы (неравенство, маргинализация населения старых районов и др.). Если не считать позитивным результатом установление хороших связей между администрацией города и компанией, проект, повлекший к отвлечению финансовых ресурсов от других более важных направлений (связанных с улучшением школьного образования и др.) оказался провальным [50].

7.6.2. Рынок решений для «умных» городов

По прогнозам компании MarketsandMarkets объем рынка решений для умных городов в мире увеличится с 410,8 млрд в 2020 до 820,7 млрд долл. в 2025 г. Предполагается, что рост рынка будет обусловлен увеличением спроса на системы безопасности граждан (видеонаблюдение, распознавание лиц и др.), увеличением численности городского населения, а также интенсификацией городских инициатив со стороны администраций. Наибольшая доля рынка будет приходиться на отрасль транспорта [7].

Значительное внимание будет уделяться разработке и внедрению систем безопасности, связанных с Интернетом вещей. Получит дальнейшее развитие инициативы в области зеленой энергетики (правительства многих стран усиливают контроль над выбросами CO₂ и других газов). Ускоренными темпами будут вводиться системы управления спросом на энергию (Demand Response Management – DRM), позволяющие регулировать пиковые электрические нагрузки [51].

Более высокими темпами возрастет такой важный сегмент рынка, как Интернет вещей: среднегодовые темпы прироста этого сегмента составят 18,1% против 14,8% всего рынка решений для умных городов. Наиболее высокими темпами

этот сегмент рынка будет развиваться в странах Восточной Азии (Китае, Японии, Сингапуре). При этом объем рынка увеличится со 113,1 млрд в 2020 до 260 млрд долл в 2025 г. Предполагается, что внедрение Интернета вещей существенно улучшит мониторинг и связь между различными сферами деятельности в городах в режиме реального времени. Однако главной проблемой для развития всего рынка и его сегмента являются чрезвычайно высокие начальные расходы на реализацию проектов. Опыт успешных проектов в ряде городов (Амстердам, Масдар в Абу-Даби и др.) свидетельствует о необходимости государственной поддержки на начальном этапе реализации проектов.

Следует отметить, что в последние годы прогнозы развития рассматриваемого рынка становятся менее оптимистичными. Так, по оценкам той же исследовательской компании Markets and Markets, объем рынка, составлявший в 2017 г. 424,68 млрд долл, должен был бы увеличиться к 2022-му до 1,2 трлн долл [52].

Работы по созданию «умных» городов ведутся и в нашей стране. С 2017 г. у нас реализуется целевая государственная программа «Цифровая экономика», в которую впоследствии был включен проект «Умный город». Планировалось, что к 2024 г. кандидатами на звание «умного города» смогут стать 18 городов из 15 регионов России. Цель проекта, реализуемого также в рамках программы «Жилье и городская среда», – «повышение конкурентоспособности российских городов, формирование эффективной системы управления городским хозяйством, создание безопасных и комфортных условий для жизни горожан...» путем широкого использования прогрессивных информационно-коммуникационных и других инженерных технологий. Проект реализуется Министерством строительства и жилищно-коммунального хозяйства Российской Федерации в рамках упомянутых выше национальных программ.

К тематике «умного» города подключились также такие крупные корпорации, как Росатом, Ростелекома, Ростех и др. а также некоторые университеты (в частности, МГУ им. М.В. Ломоносова). В конце 2017 г. некоторые из этих организаций объявили о создании Национального консорциума развития и внедрения цифровых технологий в сфере городского управления, который должен был стать центром компетенций в области создания «умных» городов. В начале 2019 г. упомянутые выше госкорпорации подписали соглашение о создании Национального центра компетенций «Умный город» [53].

В целом, объем российского рынка технологических решений для «умного города» в 2018–2019 гг. немного превышал 80 млрд руб. В 2019 г. в нашей стране было реализовано более 2,5 тыс. проектов, связанных с разработкой решений для «умных» городов, причем 40% из них составляли проекты в области автоматизации зданий,

33% – проекты цифровизации энерго- и теплоснабжения, 32% – проекты мониторинга общественной безопасности, 27% – инициативы по обращению с отходами. Вместе с тем, в исследовании консалтинговой компании iKS-Consulting отмечается незаинтересованность частных инвесторов в участии в проектах «умного города» из-за их низкой рентабельности. Характерно, что в 2018 г. 93% расходов на рынке приходилось на Москву [54].

Кроме того, по мнению экспертов из Центра стратегических разработок, наша страна отстает от стран-лидеров в создании «умных городов». В наиболее авторитетных международных рейтингах города России занимают в лучшем случае 50-е или 70-е места. Отмечается также, что среди компаний-разработчиков технологий для «умных городов» котируется лишь одна отечественная компания (Kaspersky Lab), специализирующаяся в области кибербезопасности [44].

7.6.3. Риски ИКТ в «умных» городах

С позиций специалистов в области ИКТ «умный» город можно схематично представить в виде 4-х слоев: 1) детектирующие приборы; 2) межсетевое сопряжение (шлюзы); 3) платформы (данные, искусственный интеллект, управление соединениями); 4) приложения. «Умный» город это централизованная киберфизическая система, в которой интегрировано огромное количество детекторов и других приборов, электросетей, линий мобильной связи, Интернет, объектов жизнеобеспечения и других инфраструктурных объектов, имеющих критически важное значение. Эта система характеризуется наличием сетевого эффекта: неполадки или аварии в одной из ее частей могут повлечь лавинообразный поток отказов во всей системе. Централизованные киберфизические системы, несмотря на предоставляемые ими определенные удобства и выгоды, весьма уязвимы к различного рода возмущающим воздействиям, таким, как кибер- и террористические атаки, военные действия, биотерроризм и др., направляемым против населения, бизнес-структур, производственных объектов. Более того, централизация таких систем может способствовать интенсификации различного рода преступлений против граждан «умного» города (не только экономических и связанных с утечкой персональных данных).

Учет подобных рисков и угроз тем более актуален, что в последние годы резко возросло число кибератак на энергосистемы, объекты жизнеобеспечения и критически важной инфраструктуры городов и стран. Так, по данным экспертов компании Deloitte энергетические системы США являются одними из главных целей атак хакеров, но и конкурентов. Риски нарушения работы энергосистем возрастают по мере увеличения сложности энергетических объектов и сетей, а также увеличения числа подрядчиков при их создании и обслуживании (как в связи с возможным

несоответствием характеристик вводимых разными подрядчиками приборов и программных средств, так и в связи с «человеческим фактором») [7].

В перспективе, по мере расширения масштабов распространения облачных вычислений и Интернета вещей, который стал одним из главных драйвером развития «умных» городов, следует ожидать усиления интенсивности кибератак (по оценкам компании Gartner Интернетом вещей в 2020 г. должно было быть охвачено 5,8 млрд приборов, а затраты на приборы IoT могут составить 389 млрд долл. [55]). В связи с этим, отмечается, что даже широкомасштабное использование «умных» счетчиков, устанавливаемых повсеместно быстрыми темпами, может привести к нарушениям в работе крупных энергосистем на период до нескольких месяцев [56]. Это тем более опасно, что уже в мире используется несколько сот миллионов «умных счетчиков» (в 2022 г. их число может возрасти до почти 600 млн), а ввиду «демократизации» ИКТ, любой хакер может приобрести необходимое для совершения преступления оборудование за небольшую сумму (менее 50 долл.) [57].

В связи с уязвимостью Интернета вещей, участились успешные кибератаки, в частности, DDoS-атаки, на организации и компании, использующие IoT^{xxii}. Во многом, благодаря форсированному развертыванию Интернета вещей наблюдается устойчивый рост DDoS-атак, доля которых в хакерских атаках составляла в последнее время 36% [58]. Основной задачей кибербезопасности для «умного города» является защита критически важной инфраструктуры (водо- и энергоснабжения, систем связи, данных, Интернета), а также критически важных городских служб (полиции, пожарных служб, транспорта, банков, здравоохранения и др.) [59].

Вместе с тем, должны учитываться и другие угрозы, связанные с риском утечки огромных массивов данных, циркулирующих в «умных» киберфизических системах и сетях связи (как персональных, так и общественных). Однако, как отмечается в [60], общество находится в таком же состоянии в области правовой защиты частной жизни (privacy) и персональных данных, как «во времена Дикого Запада». Весьма остро стоит и проблема информационного неравенства в «умных» городах.

Более того, отмечается, что государства, в условиях форсированного внедрения ИКТ в стремящихся «поумнеть» городах, не подготовили законодательную базу для защиты граждан от возможных негативных последствий превращения городов в «умные».

Еще один из важнейших факторов риска, связанный с медицинскими последствиями развертывания технологии мобильной связи 5G, рассмотрен в п. 7.4.2.

^{xxii} DDoS-атака (distributed denial of service – англ. распределенный отказ в обслуживании) – атака, вызывающая перегрузку на сервере или сайте жертвы путем генерации большого числа входящих запросов со стороны хакерской сети, в которой объединено большое число компьютеров.

Таким образом, рассмотренные выше нерешенные проблемы могут в среднесрочной перспективе замедлить темпы реализации проектов «умных» городов. Однако в краткосрочной перспективе, из-за ситуации, возникшей с пандемией COVID-19, высокие темпы внедрения ИКТ в городскую жизнь сохранятся.

7.7. Заключение

По мере приближения к физическим пределам кремниевой микроэлектроники, драйверами технологического развития производителей элементной базы становятся такие области применения, как центры обработки данных (ЦОД), мобильные устройства, Интернет вещей (IoT) и различные приложения искусственного интеллекта (ИИ).

В связи с ожидаемым позитивным влиянием ИКТ и, особенно ИИ, на развитие бизнеса и экономики, а также стремлением ряда государств к усилению своего военного потенциала, в настоящее время разворачивается острая глобальная конкуренция в области ИКТ и ИИ.

Проведенное исследование свидетельствует о необходимости учета большого числа факторов риска при оценке перспектив ИКТ. Хотя значительная часть этих факторов связана с господствующими в экономике законами рынка, ориентирующие бизнес на быструю прибыль и обогащение, решающую роль среди них играют этические факторы.

С целью снижения риска негативного влияния ИКТ на общественное развитие, необходимо уже в самое ближайшее время разработать нормативно-законодательную базу и кодекс этических норм и правил, которыми должны руководствоваться разработчики, заказчики и пользователи ИКТ-приборов и услуг. Сертификацию продуктов и технологий ИКТ следует проводить только с учетом оценки социально-экономических рисков их использования.

Необходимо расширить фронт работ научно-исследовательских организаций в области системного анализа и оценки социально-экономических рисков и перспектив использования ИКТ, а также обеспечить участие этих организаций в проведении экспертиз проектов ИКТ.

Необходимо обеспечивать привлечение общественности к широкому демократическому очному (невиртуальному) обсуждению перспектив использования новых направлений ИКТ, в частности, проведение общественных экспертиз, референдумов и др.

Концепция «умного» города возникла под влиянием бурного развития процессов урбанизации, а также информационно-коммуникационных технологий (ИКТ).

Имеет место несопоставимость рейтингов различных организаций, используемых для оценки прогресса в реализации проектов создания «умных» городов.

Прогнозы развития рынка решений для «умных» городов становятся все менее оптимистичными, что связано с необходимостью осуществления огромных начальных затрат на реализацию проектов, с нерешенностью ряда остро стоящих проблем безопасности, социальных и медицинских проблем, а также с отсутствием законодательной базы для защиты граждан от возможных негативных последствий превращения городов в «умные».

Список литературы к главе 7

1. Дубинина М.Г. Моделирование динамики взаимосвязи макроэкономических показателей и показателей распространения ИТ в развитых и развивающихся странах // Труды ИСА РАН. 2015. № 65-1. С. 24–37.
2. URL: <https://spectrum.ieee.org/tech-talk/robotics/artificial-intelligence/former-head-of-google-china-foresees-an-ai-crisis/> (дата доступа 05.09.2018).
3. Варшавский Л.Е. Кризис финансовой системы и эволюция товарных рынков // Прикладная эконометрика. 2010. № 1 (17). С. 30–44.
4. Дертузос М.Л. Передача данных, компьютеры и сети // В мире науки. 1991. № 11. С. 6–17.
5. Маршалл Дж.А., Бансал В.К. Финансовая инженерия: пер. с англ. М.: ИНФРА-М, 1998. 784 с.
6. Варшавский Л.Е. Анализ социально-экономических проблем и вызовов, связанных с ИКТ// Труды Института системного анализа Российской академии наук. 2019. Т. 69. Вып. 1. С. 3–16.
7. Варшавский Л.Е. Концепция «умного города» и риски информационно-коммуникационных технологий (ИКТ) // Теория и практика институциональных преобразований в России: сб. научных трудов / Ерзнкян Б.Г. (ред.). Вып. 50. М.: ЦЭМИ РАН, 2020. С. 116–127.
8. Modeling Artificial Intelligence and Exploring its Impact. Pardee Center for International Futures. Working Paper. May 2017.
9. The Impact of Artificial Intelligence on Innovation. I. M. Cockburn, R. Henderson, and S. Stern. NBER Working Paper No. 24449. March 2018.
10. Factories of the Future Need AI to Survive and Compete. URL: <https://www.forbes.com/sites/mariyayaao/2017/08/08/industrial-ai-factories-of-future/#79e426d3128e> (дата доступа 25.08.2018).
11. Feldman M. DARPA Sets Up Fast Track for Third Wave AI. URL: <https://www.top500.org/news/darpa-sets-up-fast-track-for-third-wave-ai/> (дата доступа 21.07.2018).
12. Charles A.S. Interpreting Deep Learning: The Machine Learning Rorschach Test? // Siam News. 2018. July/August.
13. Варшавский Л.Е. Экономико-математическое моделирование тенденций изменения ключевых показателей высокопроизводительных суперкомпьютеров // Труды Института системного анализа Российской академии наук. 2017. Т. 67. № 3. С. 64–74.
14. Feldman M. The DOE and NSA Construct Doomsday Scenario for American HPC. URL: <https://www.top500.org/news/the-doe-and-nsa-construct-doomsday-scenario-for-american-hpc/> (дата доступа 07.04.2017).
15. Preparing for the Future of Artificial Intelligence. National Science and Technology Council. December 2016. Executive Office of the President. Washington, D.C. 20502. December 20, 2016.
16. China Science&Technology Newsletter, No 17, September 15, 2017.

17. Acemoglu D., Restrepo P. Artificial Intelligence, Automation and Work // Journal of Human Capital. 2018. Vol. 12. No. 2.
18. Bughin J., Hazan E., Lund S., Dahlström P., Wiesinger A., Subramaniam A. Skill shift: Automation and the future of the workforce. Discussion Paper. McKinsey Global Institute, May 2018. URL: <https://www.mckinsey.com/featured-insights/future-of-work/skill-shift-automation-and-the-future-of-the-workforce> (дата доступа 30.08.2018).
19. AI taking jobs is not a concern for certain workers. URL: https://searchbusinessanalytics.techtarget.com/tip/3-ways-to-make-machine-learning-in-business-more-effective?track=NL-1816&ad=923710&src=923710&asrc=EM_NLN_102097887&utm_medium=EM&utm_source=NLN&utm_campaign=20181018_How%20to%20do%20machine%20learning%20the%20right%20way%20for%20business%20uses;%20jobs%20that%20AI%20isn%27t%20likely%20to%20take;%20Oracle%20database%20performance%20management%20tips (дата доступа 15.08.2018).
20. Маркс К. Капитал. Критика политической экономии. Т. 1. Кн. 1. М.: Политиздат, 1978.
21. The Future of Jobs Report. 2018. World Economic Forum. Centre for the New Economy and Society.
22. Винер Н. Творец и робот. Гл. V. М.: Прогресс, 1966.
23. Бир С. Мы и сложность современного мира // Кибернетика сегодня: проблемы и суждения: сборник. М.: Знание, 1976. С. 7–39.
24. Визнер Дж.Б. Прспективы информационной тирании // Кибернетика сегодня: проблемы и суждения: сборник. М.: Знание, 1976. С. 39–46.
25. Проблемы, возникающие в связи с взаимоотношениями ЭВМ и общества // Кибернетика сегодня: проблемы и суждения: сборник. М.: Знание, 1976. С. 46–62.
26. Брукшир Дж. Информатика и вычислительная техника. М.: Питер, 2004. 618 с.
27. Cukier K., Mayer-Schoenberger V. The Rise of Big Data. How It's Changing the Way We Think About World // Foreign Affairs. 2013. May/June. Pp. 28–40.
28. Иванов В.В., Малинецкий Г.Г. Цифровая экономика: мифы, реальность, перспектива. М.: Российская академия наук, 2017.
29. Шахрай С.М. Цифровая конституция. Основные права и свободы личности в тотально информационном обществе // Вестник Российской академии наук. 2018. Т. 88. № 12. С. 1075–1082.
30. Под колпаком у PRISM. Раскрыта гигантская секретная программа слежки американских спецслужб за пользователями интернета // Взгляд.ру. 2013. 7 июня.
31. Уитэкер Л. Как навсегда удалиться из соцсетей // Метро. 2018. 29 марта. С. 08.
32. URL: <https://www.radiationhealthrisks.com/5g-cell-towers-dangerous/>
33. URL: <https://www.scientificamerican.com/article/new-studies-link-cell-phone-radiation-with-cancer/>
34. Варшавский А.Е. Проблемные инновации: риски для человечества. Экономические, социальные и этические аспекты. М.: ЛЕНАНД, 2014.
35. Захаров А. Тотальная слежка: как устроен рынок торговли пользовательскими данными. URL: https://www.rbc.ru/magazine/2018/04/5aafdfc99a7947654297214d?from=center_2
36. Фейнберг А. Каждого второго заменит робот. URL: <https://www.rbc.ru/newspaper/2018/03/30/5abcc7729a7947e437dc77a7>
37. Жидких И., Серебрянный В. Безлюдные производства: как роботизировать российскую экономику. URL: https://www.rbc.ru/opinions/economics/06/03/2018/5a9e49c39a79473e0ef8c42b?from=materials_on_subject
38. URL: <https://unhabitat.org/global-state-of-metropolis-2020-%E2%80%9393-population-data-booklet> (дата доступа 25.08.2020).
39. Bouchet M., Sifan Liu, Parilla J., Kabbani N. Global Metro Monitor. Metropolitan Policy Program at Brookings, 2018.

40. Рифкин Дж. Третья промышленная революция: как горизонтальные взаимодействия меняют энергетику, экономику и мир в целом / пер. с англ. М.: Альпина нон-фикшн, 2014. 410 с.
41. Smart Cities – A Deloitte Point of View, Version 1.0, November 2015.
42. URL: <http://www.itu.int/en/ITU-T/focusgroups/ssc/Pages/default.aspx> (дата доступа 25.11.2019).
43. Ericsson networked society city index 2016.
44. Приоритетные направления внедрения технологий умного города в российских городах. Центр стратегических разработок. М., 2018.
45. URL: http://smartcity.cnews.ru/articles/2018-07-30_moskva_stala_polnopravnym_uchastnikom_mezhdunarodnyh_rejtingov (дата доступа 25.11.2019).
46. Silvaa B.N., Khanb M., Hana R. Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities // Sustainable Cities and Society. 2018. No. 38. Pp. 697–713.
47. Smart City Index 2020. IMD-SUTD Smart City Index Report.
48. Engelbert J, van Zoonen L., Hirzalla F. Excluding citizens from the European smart city; The discouse practices of pursuing and granting smartness // Technological Forecasting & Social Change. 2019. No. 142. Pp. 347–353.
49. Martin C.J., Evans J., Karvonen A. Smart and sustainable? Five tensions in the visions and practices of the smart-sustainable city in Europe and North America // Technological Forecasting & Social Change. 2018. No. 133. Pp. 269–278.
50. URL: https://works.bepress.com/alan_wiig/2/ (дата доступа 23.10.2019).
51. URL: https://www.marketsandmarkets.com/Market-Reports/smart-cities-market-542.html?gclid=EAIaIQobChMIg8C_mK6M7AIVywJ7Ch06XAn-EAAYASAAEgLuXfD_BwE (дата доступа 25.11.2019).
52. Smart Cities Market worth 1,201.69 Billion USD by 2022 // Markets and Markets. URL: <https://www.marketsandmarkets.com/PressReleases/smart-cities.asp> (дата доступа: 19.03.2018).
53. URL: <https://rostec.ru/news/umnyy-gorod-stolitsa-tsifrovoy-ekonomiki/> (дата доступа 25.11.2019).
54. URL: https://www.tadviser.ru/index.php/%D0%A1%D1%82%D0%B0%D1%82%D1%8C%D1%8F:%D0%A3%D0%BC%D0%BD%D1%8B%D0%B9_%D0%B3%D0%BE%D1%80%D0%BE%D0%B4:_%D1%80%D0%B0%D0%B7%D0%B2%D0%B8%D1%82%D0%B8%D0%B5_%D0%B2_%D0%A0%D0%BE%D1%81%D1%81%D0%B8%D0%B8 (дата доступа 25.05.2020).
55. URL: <https://www.gartner.com/en/newsroom/press-releases/2019-08-29-gartner-says-5-8-billion-enterprise-and-automotive-iiot> (Доступ 11.09.2020).
56. URL: <https://phys.org/news/2017-08-smart-electrical-grids-vulnerable-cyber.html> (дата доступа 15.11.2019).
57. URL: <https://techxplore.com/news/2019-06-ways-hackproof-smart-meters.html> (дата доступа 25.01.2020).
58. URL: <https://securityboulevard.com/2020/06/survey-finds-sluggish-cybersecurity-response-to-pandemic/> (дата доступа 25.07.2020).
59. Chai K. Toh. Security for smart cities // IET Smart Cities. 2020. Vol. 2. Iss. 2. Pp. 95–104.
60. URL: <https://www.comparitech.com/blog/vpn-privacy/smart-cities-privacy-risks/> (дата доступа 29.09.2020).

Монография

Л.Е. Варшавский

СОЦИАЛЬНО-ЭКОНОМИЧЕСКИЕ
ПРОБЛЕМЫ РАЗВИТИЯ
ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫХ
ТЕХНОЛОГИЙ (ИКТ)

Подписано в печать 22.08.2022 г.

Формат 60×90/16. Печ. л. 10,0. Тираж 100 экз. Заказ № 5

ФГБУН Центральный экономико-математический институт РАН

117418, Москва, Нахимовский пр., 47

Тел. 8 (499) 724-21-39

E-mail: ecr@cemi.rssi.ru

<http://www.cemi.rssi.ru/>

ISBN 978-5-8211-0805-0



9 785821 108050